

**PERINGKASAN ARTIKEL BERBAHASA INDONESIA
MENGUNAKAN *TEXTRANK* DENGAN PEMBOBOTAN BM25**

SKRIPSI

Disusun oleh:
Yurdha Fadhila Hernawan
NIM: 165150200111094



PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BRAWIJAYA
MALANG
2020

PERSETUJUAN

PERINGKASAN ARTIKEL BERBAHASA INDONESIA MENGGUNAKAN TEXTRANK
DENGAN PEMBOBOTAN BM25

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun Oleh :
Yurdha Fadhila Hernawan
NIM: 165150200111094

Telah diperiksa dan disetujui oleh:

Dosen Pembimbing I

Dosen Pembimbing 2

Putra Pandu Adikara, S.Kom., M.Kom.
NIK: 19850725 200812 1 002

Randy Cahya Wihandika, S.ST., M.Kom.
NIK: 201405 880206 1 001

Mengetahui
Ketua Jurusan Teknik Informatika

Tri Astoto Kurniawan, S.T, M.T., Ph.D
NIP: 19710518 200312 1 001

KATA PENGANTAR

Untuk banyaknya kesempatan yang berhasil diambil sejauh ini, penulis mengucapkan terima kasih atas Tuhan Yang Maha Esa sehingga dapat menyelesaikan tugas akhir jenjang perkuliahan Strata-1 dengan judul “Peringkasan Artikel Berbahasa Indonesia Menggunakan *TextRank* dengan Pembobotan BM25”. Perjalanan sejauh ini tentu saja tidak luput dari kesalahan yang pernah penulis ambil dan banyaknya dukungan serta doa dari pihak-pihak terdekat. Oleh karena itu, penulis ingin menyampaikan rasa hormat dan terima kasih kepada:

1. Mami, Papi, Amel, Fadhil dan seluruh keluarga atas kasih sayang, pengertian dan kesabarannya untuk selalu membimbing dan menuntun penulis menjadi pribadi yang lebih baik.
2. Bapak Putra Pandu Adikara, S.Kom., M.Kom. selaku dosen pembimbing 1 yang telah membimbing dan membantu setiap langkah dari penulisan skripsi ini.
3. Bapak Randy Cahya Wihandika, S.ST., M.Kom. selaku dosen pembimbing 2 yang telah memberikan banyak masukan demi lancarnya penulisan skripsi ini.
4. Seluruh dosen, guru, dan civitas akademik Teknik Informatika Fakultas Ilmu Komputer Universitas Brawijaya yang telah memberikan banyak pelajaran dan ilmu yang bermanfaat.
5. Keluarga besar EMIF 2019/2020 yang selalu memberikan tempat terbaik selama proses perkuliahan.
6. Orang-orang terdekat; Nisrina, Farah, Benita, Citra, Titus, Allysa, Fun, Ilham, Nurul, Noraesa, dan Yahya. Terima kasih karna selalu ada.
7. Kopi Lain Hati Panjaitan dan Spotify yang telah menemani sepanjang penulisan dari skripsi ini.

Banyaknya kesalahan yang mungkin ditemukan pada laporan ini sangat membantu penulis dalam memperbaiki diri dikemudian hari, untuk itu penulis sangat terbuka apabila ada kritik dan saran yang ingin disampaikan. Semoga tugas akhir ini dapat berguna dan membantu semua pihak yang membacanya.

Malang, 29 Februari 2020

Yurdha Fadhila Hernawan
yurdhafadhila@gmail.com

ABSTRAK

Yurdha Fadhila Henawan, Peringkasan Artikel Berbahasa Indonesia Menggunakan *Textrank* Dengan Pembobotan BM25

Pembimbing : Putra Pandu Adikara, S.Kom., M.Kom. dan Randy Cahya Wihandika, S.ST., M.Kom.

Penggunaan internet sebagai sumber informasi telah membawa manusia pada era *one click away*. Apa pun bisa diakses di mana pun kapan pun, baik secara visual maupun tidak. Namun, tidak semua informasi yang diakses selalu sesuai dengan konteks yang diinginkan. Untuk memudahkan pengguna internet dalam mendapatkan informasi yang ringkas dengan tidak merusak atau menghilangkan informasi penting, maka dibutuhkan suatu peringkasan otomatis. Salah satu cara untuk mendapatkan ringkasan pada sebuah dokumen adalah dengan mencari kumpulan kalimat penting pada dokumen yang dapat merepresentasikan dokumen asli secara keseluruhan. Metode peringkasan tersebut disebut juga dengan peringkasan ekstraktif. Pada penelitian ini, peringkasan ekstraktif dilakukan dengan memeringkatkan setiap kalimat pada sebuah dokumen dan mengambil kalimat dengan peringkat teratas sebagai ringkasan. Metode *TextRank* yang digunakan pada penelitian ini merepresentasikan dokumen sebagai graf, setiap kalimat dianggap sebagai *node* dan hubungan antara kalimat (*node*) merupakan nilai *similarity* antar kalimat. Fungsi *similarity* yang digunakan adalah BM25 dengan metode pemeringkatan *PageRank*. Panjang ringkasan yang dihasilkan sistem disesuaikan dengan besar nilai *compression rate* yang digunakan. Setelah membandingkan hasil ringkasan yang didapatkan sistem peringkasan otomatis dengan hasil ringkasan yang didapatkan dari *expert* (pakar) sebanyak 10 dokumen, kualitas ringkasan terbaik didapatkan pada saat penggunaan *compression rate* sebesar 30% dengan nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,551692; 0,551692; dan 0,551692.

Kata kunci : peringkasan otomatis, *TextRank*, BM25, *PageRank*, *compression rate*

ABSTRACT

Yurdha Fadhila Henawan, Peringkasan Artikel Berbahasa Indonesia Menggunakan *Textrank* Dengan Pembobotan BM25

Pembimbing : Putra Pandu Adikara, S.Kom., M.Kom. dan Randy Cahya Wihandika, S.ST., M.Kom.

The use of the internet as a source of information has brought humans to a one-click era. Anything can be accessed anywhere, visually or not. However, not all informations accessed always match with the context itself. An automatic summarization is needed to help people to get the concise informations without ruin the context and missing the point. One way to get a summarize of the document is to find a collection of important sentences in the document that can represent the original document as a whole. That automatic text summarization method is also called extractive summarize. In this study, extractive summarization is done by checking each sentence in a document and ranking the important sentences. The TextRank method used in this study will represent the document as a graph, each sentence will be considered as a node and the relationship between sentences (nodes) is the value of similarity between sentences. The similarity function used is BM25 with the PageRank as ranking method. The resulting length of the system will be adjusted to the value of the level of compression used. After comparing the summarization result between the automatic system and an expert of 10 documents, the best quality is obtained when using a compression rate of 30% with an average value of precision, recall, and f-measure is 0.551692; 0.551692; and 0.551692.

Keywords : automatic summarization, *TextRank*, BM25, *PageRank*, *compression rate*

DAFTAR ISI

PERSETUJUAN	i
KATA PENGANTAR.....	ii
ABSTRAK.....	iv
ABSTRACT	v
DAFTAR ISI	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan	2
1.4 Manfaat.....	2
1.5 Batasan Masalah	2
1.6 Sistematika Pembahasan	3
BAB 2 LANDASAN KEPUSTAKAAN	4
2.1 Kajian Pustaka	4
2.2 Dasar Teori Peringkasan Teks	6
2.3 <i>Web Scrapping</i>	7
2.4 <i>Preprocessing</i>	8
2.4.1 Segmentasi	8
2.4.2 <i>Cleaning</i>	8
2.4.3 Stemming	8
2.4.4 Tokenisasi	8
2.5 <i>Term Weighting</i>	9
2.5.1 <i>Term Frequency (tf)</i>	9
2.5.2 <i>Document Frequency (df)</i>	9
2.5.3 <i>Inverse Document Frequency (idf)</i>	9
2.6 <i>TextRank</i>	10
2.7 Fungsi <i>Similarity</i> BM25	10
2.8 <i>PageRank</i>	11

2.9 <i>Compression Rate</i>	13
2.10 Evaluasi	13
BAB 3 METODOLOGI	15
3.1 Jenis Penelitian	15
3.2 Pengambilan Data	15
3.3 Metode Penelitian	15
3.4 Peralatan Pendukung.....	16
3.5 Teknik Analisis Data	16
BAB 4 PERANCANGAN.....	18
4.1 Perancangan Sistem.....	18
4.1.1 Perancangan Umum Sistem Peringkasan Otomatis	18
4.1.2 <i>Web Scrapping</i>	19
4.1.3 Perancangan <i>Preprocessing</i>	19
4.1.4 Perancangan <i>Term Weighting</i>	21
4.1.5 Perancangan Perhitungan BM25	22
4.1.6 Perancangan Perhitungan <i>PageRank</i>	23
4.1.7 Perancangan Mendapatkan Ringkasan.....	24
4.2 Perhitungan Manual	24
4.2.1 Data Perhitungan Manual	25
4.2.2 <i>Preprocessing</i>	25
4.2.3 <i>Term Weighting</i>	29
4.2.4 Perhitungan BM25	32
4.2.5 Perhitungan <i>PageRank</i>	35
4.2.6 Mendapatkan Ringkasan.....	38
4.3 Perancangan Evaluasi	39
BAB 5 IMPLEMENTASI	41
5.1 Implementasi <i>Web Scrapping</i>	41
5.2 Implementasi <i>Preprocessing</i>	42
5.3 Implementasi <i>Term Weighting</i>	44
5.4 Implementasi Kalimat Sebagai <i>Node</i>	46
5.5 Implementasi Graf Dokumen.....	47
5.6 Implementasi <i>Main</i> Sistem Peringkasan	50

5.7 Hasil Implementasi	52
BAB 6 PENGUJIAN DAN ANALISIS.....	53
6.1 Hasil Pengujian dan Analisis <i>Compression Rate</i> 5%.....	53
6.2 Hasil Pengujian dan Analisis <i>Compression Rate</i> 10%.....	54
6.3 Hasil Pengujian dan Analisis <i>Compression Rate</i> 20%.....	55
6.4 Hasil Pengujian dan Analisis <i>Compression Rate</i> 30%	56
BAB 7 PENUTUP	59
7.1 Kesimpulan.....	59
7.2 Saran	59
DAFTAR RUJUKAN	60
LAMPIRAN A SURAT PERNYATAAN PARTISIPASI PAKAR.....	62
LAMPIRAN B HASIL KOMPARASI RINGKASAN SISTEM DAN RINGKASAN PAKAR.....	63

DAFTAR TABEL

Tabel 2.1 Contoh Hasil Segmentasi Dokumen	8
Tabel 2.2 Contoh Hasil Stemming Dokumen	8
Tabel 2.3 Contoh Hasil Perhitungan Tf	9
Tabel 2.4 Contoh Hasil Perhitungan Df dan Idf	10
Tabel 2.5 Contoh <i>Confusion Matrix</i>	13
Tabel 4.1 Hasil Segmentasi.....	25
Tabel 4.2 Hasil <i>Cleaning</i>	27
Tabel 4.3 Hasil Stemming.....	28
Tabel 4.4 Hasil Tokenisasi	29
Tabel 4.5 Hasil Perhitungan tf	30
Tabel 4.6 Hasil Perhitungan df	30
Tabel 4.7 Hasil Perhitungan idf	31
Tabel 4.8 Hasil Perhitungan BM25	34
Tabel 4.9 Inisialisasi Awal Bobot <i>PageRank</i>	35
Tabel 4.10 Hasil Perhitungan Bobot <i>PageRank</i>	38
Tabel 4.11 Hasil Pengurutan Bobot <i>PageRank</i>	38
Tabel 4.12 Hasil Ringkasan Dokumen	39
Tabel 4.13 Perancangan Evaluasi Sistem Compression Rate 5%.....	39
Tabel 4.14 Perancangan Evaluasi Sistem Compression Rate 10%.....	40
Tabel 4.15 Perancangan Evaluasi Sistem Compression Rate 20%.....	40
Tabel 4.16 Perancangan Evaluasi Sistem Compression Rate 30%.....	40
Tabel 6.1 Hasil Pengujian <i>Compression Rate</i> 5%	53
Tabel 6.2 Hasil Pengujian <i>Compression Rate</i> 10%	54
Tabel 6.3 Hasil Pengujian <i>Compression Rate</i> 20%	55
Tabel 6.5 Hasil Pengujian <i>Compression Rate</i> 30%	57

DAFTAR GAMBAR

Gambar 2.1 Contoh Hasil Tokenisasi Kalimat Index 1	9
Gambar 2.2 Struktur <i>Textrank</i> Sebagai Graf	11
Gambar 2.3 Struktur <i>Textrank</i> Setelah Perhitungan <i>Pagerank</i>	12
Gambar 3.1 Gambaran Umum Proses Peringkasan <i>Single Document</i>	16
Gambar 4.1 Diagram Alir Sistem	19
Gambar 4.2 Diagram Alir <i>Web Scrapping</i>	20
Gambar 4.3 Diagram Alir <i>Preprocessing</i>	20
Gambar 4.4 Diagram Alir <i>Term Weighting</i>	22
Gambar 4.5 Diagram Alir Perhitungan BM25	23
Gambar 4.6 Diagram Alir Perhitungan <i>PageRank</i>	24
Gambar 4.7 Diagram Alir Pembentukan Ringkasan	24
Gambar 4.8 Dokumen Perhitungan Manualisasi	25
Gambar 5.1 Contoh Hasil Peringkasan Sistem	52
Gambar 6.1 Evaluasi Pengujian <i>Compression Rate</i> 5%	54
Gambar 6.2 Evaluasi Pengujian <i>Compression Rate</i> 10%	55
Gambar 6.3 Evaluasi Pengujian <i>Compression Rate</i> 20%	56
Gambar 6.4 Contoh Hasil Kalimat Yang Terpilih Sebagai Ringkasan	57
Gambar 6.5 Evaluasi Pengujian <i>Compression Rate</i> 30%	58

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Penggunaan internet sebagai sumber informasi telah membawa manusia pada era *one click away*. Apa pun bisa diakses di mana pun kapan pun, baik secara visual maupun tidak. Namun, tidak semua informasi yang diakses selalu sesuai dengan konteks yang diinginkan. Bisa dikatakan hanya sedikit pengguna yang dapat memahami semua informasi ketika membaca sebuah tulisan panjang (Niu et al., 2016). Kesulitan tersebut dapat membuat pengguna untuk membaca ulang, sehingga menghabiskan banyak waktu.

Untuk memudahkan pengguna internet dalam mendapatkan informasi yang ringkas dengan tidak merusak atau menghilangkan informasi penting, maka dibutuhkan suatu peringkasan otomatis (Abbasi-ghalehtaki et al., 2016). Berdasarkan metode yang digunakan, peringkasan teks dapat dikategorikan dalam dua bentuk, yaitu ekstraktif dan abstraktif. Ringkasan ekstraktif merupakan ringkasan yang terdiri atas kumpulan dari bagian-bagian penting suatu tulisan yang dapat mewakili keseluruhan teks, sedangkan ringkasan abstraktif merupakan ringkasan yang terdiri dari kalimat baru yang dapat merepresentasikan konteks tulisan dalam bentuk lain. Selain itu, peringkasan teks juga dapat dikelompokkan berdasarkan jumlah dokumen yang digunakan menjadi *single document* dan *multi-document* (Fang et al., 2017).

Masalah utama yang muncul setelah melakukan peringkasan adalah kualitas hasil peringkasan. Apakah konteks yang dibicarakan pada hasil ringkasan sudah dapat merepresentasikan tulisan secara utuh. Penelitian oleh Pinandhita (2013) telah melakukan peringkasan ekstraktif teks berbahasa Indonesia dengan menggunakan kata benda yang terdapat dalam sebuah dokumen. Penelitian tersebut menghitung nilai kemiripan (*similarity*) antara kalimat berdasarkan kata benda yang terdapat pada setiap kalimat, lalu memeringkatkan kalimat yang paling penting dengan mengurutkan total nilai kemiripan tersebut. Evaluasi dilakukan dengan membandingkan hasil kinerja beberapa metode *similarity* agar dapat menentukan metode *similarity* terbaik dalam melakukan peringkasan. Pada penelitian tersebut didapatkan hasil bahwa penggunaan metode *similarity* BM25 menghasilkan ringkasan yang lebih baik dari pada metode *similarity* lainnya.

Selain itu, peneliti lain telah melakukan penelitian untuk mendapatkan ringkasan secara abstraktif (Niu et al., 2016). Penelitian tersebut menggunakan dokumen teks opini pendek berbahasa China dengan mengelompokkan teks yang mirip lalu meringkasnya. Proses pengelompokan teks yang mirip dilakukan dengan menggunakan metode *K-Means* dengan fitur yang didapatkan dari nilai *word2vec* (nilai hubungan kedekatan antara kata). Hasil pengelompokan teks diperingkatkan menggunakan metode *TextRank*. Kalimat yang dianggap penting dan dapat dijadikan ringkasan memiliki peringkat tertinggi. Hasil pemeringkatan dijadikan

ringkasan dengan menggunakan *encoder-decoder Recurrent Neural Network (RNN)* untuk membentuk kalimat baru.

Penelitian sebelumnya yang menjadi acuan dalam penelitian ini adalah peringkasan ekstraktif menggunakan data *statistical* (Mussina, Aubakirov and Trigo, 2018). Data *statistical* merupakan nilai kesamaan *text units* dalam sebuah dokumen. *Text units* dalam dokumen dapat berupa kata, kalimat atau paragraf. Penelitian tersebut menggunakan teks berbahasa Rusia dan Kazakhstan karena memiliki struktur kalimat yang jelas. Penelitian tersebut menggunakan metode *TextRank* untuk merepresentasikan kalimat dalam sebuah dokumen dan menggunakan beberapa metode *similarity*.

Berdasarkan pemaparan di atas, penulis mengajukan penelitian untuk melakukan peringkasan otomatis dengan objek artikel berita *online* berbahasa Indonesia menggunakan metode pemeringkatan *TextRank* dan BM25 sebagai fungsi *similarity*. BM25 dipilih berdasarkan penelitian sebelumnya yang telah melakukan pengujian menggunakan beberapa fungsi *similarity* dan mendapatkan bahwa BM25 menghasilkan ringkasan yang lebih baik dari pada fungsi *similarity* lainnya (Barrios et al., 2016).

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disebutkan, dapat dirumuskan masalah sebagai berikut:

1. Bagaimana menghasilkan ringkasan ekstraktif dokumen berbahasa Indonesia menggunakan metode *TextRank* dengan pembobotan BM25?
2. Bagaimana pengaruh *compression rate* dengan kualitas ringkasan?

1.3 Tujuan

1. Menghasilkan ringkasan ekstraktif dokumen berbahasa Indonesia menggunakan metode *TextRank* dengan pembobotan BM25
2. Membandingkan kualitas ringkasan berdasarkan *compression rate*

1.4 Manfaat

Selain menghasilkan ringkasan secara otomatis, penelitian ini juga dapat digunakan dalam *Question Answering System*, *Information retrieval*, dan ekstraksi informasi.

1.5 Batasan Masalah

1. Artikel online yang digunakan adalah artikel berita berbahasa Indonesia yang memiliki tulisan sesuai Ejaan Bahasa Indonesia (EBI)
2. Artikel online berasal dari situs resmi BBC Indonesia dengan topik bencana alam dan lingkungan sebanyak 10 artikel.

1.6 Sistematika Pembahasan

Laporan penelitian ini terdiri atas beberapa bagian yaitu sebagai berikut.

Bab I Pendahuluan

Berisikan latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat, batasan masalah yang diberikan, dan sistematika pembahasan.

Bab II Landasan Kepustakaan

Berisikan pembahasan mengenai teori yang dijadikan pedoman pengerjaan, konsep pengerjaan penelitian, metode atau sistem yang diterapkan, pustaka ilmiah yang berkaitan dengan *Text Summarize*, dan *TextRank*.

Bab III Metode Penelitian

Berisikan langkah pengerjaan dalam penelitian, teknik yang digunakan, data yang digunakan, dan representasi berdasarkan metode yang dipilih untuk menyelesaikan masalah.

Bab IV Perancangan

Berisikan proses penyelesaian masalah dengan visualisasi diagram alir, perhitungan manual dari metode yang digunakan, dan rancangan pengujian.

Bab V Implementasi

Berisikan implementasi metode yang digunakan dalam menyelesaikan masalah dalam sebuah sistem.

Bab VI Pengujian dan Analisis

Berisikan scenario pengujian dan analisis terhadap hasil pengujian.

Bab VII Penutup

Berisikan kesimpulan dan saran sebagai rekomendasi untuk penelitian selanjutnya.

BAB 2 LANDASAN KEPUSTAKAAN

2.1 Kajian Pustaka

Salah satu penelitian yang dilakukan untuk melakukan peringkasan adalah dengan memanfaatkan kata benda yang terdapat pada sebuah dokumen (Pinandhita, 2013). Penelitian tersebut dilakukan untuk mendapatkan ringkasan ekstraktif dengan membandingkan beberapa metode *similarity* yang didapatkan dari kata benda. Dokumen yang digunakan berupa artikel koran dengan topik di luar pertanian yang juga digunakan pada penelitian Miptahudin (2010) dan dokumen dari penelitian Aristoteles (2011). Hasil peringkasan yang didapatkan dari beberapa metode *similarity* dievaluasi untuk mendapatkan metode *similarity* terbaik dalam melakukan peringkasan.

Total percobaan yang dilakukan Pinandhita (2013) adalah sebanyak tujuh kali (dengan mengikutsertakan judul dokumen dan tanpa judul dokumen) yaitu penerapan *PageRank* dengan bobot *cosine similarity*, *PageRank* dengan bobot BM25, *PageRank* dengan bobot *content overlap*, bobot *cosine similarity* tanpa *PageRank*, bobot BM25 tanpa *PageRank*, bobot *content overlap* tanpa *PageRank*, dan yang terakhir menggunakan bobot koefisien *dice*. Rata rata panjang dokumen yang digunakan pada penelitian tersebut adalah sebanyak 47 kalimat yang mengandung rata rata sebanyak 282 kata benda. Pengujian dilakukan dengan mencari nilai *kappa* berdasarkan kesepakatan hasil ringkasan antara tiga orang dosen Jurusan Sastra Indonesia dan tiga orang mahasiswa. Kualitas ringkasan terbaik didapatkan dari hasil pembobotan BM25 dengan *PageRank* tanpa menggunakan judul dokumen.

Selain mendapatkan ringkasan secara ekstraktif, peneliti lain juga telah melakukan penelitian untuk mendapatkan ringkasan secara abstraktif (Niu et al., 2016). Penelitian tersebut menggunakan dokumen teks opini pendek berbahasa China dengan mengelompokkan teks yang mirip lalu meringkasnya. Proses peringkasan dibagi atas tiga tahap yaitu *clustering* teks (mengelompokkan), pemeringkatan teks, dan peringkasan teks. Proses *clustering* dilakukan dengan menggunakan metode *K-Means* dengan fitur yang telah didapatkan dari nilai *word2vec*. *Word2vec* menunjukkan hubungan kedekatan antara suatu kata dengan kata lainnya.

Hasil *clustering* yang telah didapatkan diperingkatkan menggunakan model pemeringkatan graf yaitu *TextRank* dengan bantuan pembobotan *similarity* BM25, setelahnya hasil pemeringkatan tertinggi diringkas secara abstraktif. Peringkasan abstraktif dilakukan dengan metode *encoder-decoder Recurrent Neural Network (RNN)*. Pengujian kualitas ringkasan yang dihasilkan pada penelitian Niu et al., (2016) dilakukan dengan mencari nilai *precision*, *recall*, *F-measure*, *ROUGE-N*, dan *ROUGE-L*. Secara berurut rata-rata nilai *precision*, *recall*, dan *F-measure* yang didapatkan adalah sebesar 0,94; 0,932; dan 0,933. Nilai *ROUGE-N* didapatkan dari hasil kesamaan *N-gram* yang serupa dari hasil peringkasan manual dengan hasil

ringkasan yang didapatkan sistem. Pada evaluasi ini digunakan ROUGE-1 (*unigrams*) dan ROUGE-2 (*bigrams*).

Peringkasan dokumen juga dapat dilakukan dengan metode klasifikasi kalimat (Fhadli, 2017). Kalimat diklasifikasikan sebagai kelas yang termasuk ringkasan dan kelas yang termasuk bukan ringkasan. Fitur yang digunakan dalam pengklasifikasian tersebut adalah fitur statistik dan fitur linguistik, sedangkan dokumen yang digunakan dalam penelitian tersebut adalah literatur Ilmu Komputer berbahasa Indonesia.

Fitur statistik didapatkan dari nilai TF-IDF sebuah kalimat, sedangkan fitur linguistik didapatkan dari banyaknya kata pada judul yang terdapat pada kalimat (*title word*), posisi kalimat pada dokumen (*sentence location*), panjang kalimat (*sentence length*), kalimat yang mengandung akronim atau singkatan (*upper-case word*), kalimat yang memiliki frasa yang dianggap penting seperti "jadi", "hasilnya", dan "kesimpulannya" (*cue phrase*), kalimat yang mengandung kata spesifik yang menyatakan sebuah topik dokumen (*biased word*), dan kalimat yang mengandung kata-kata yang dianggap tidak penting (*occurrence of non-essential information*). Proses pengklasifikasian kalimat dilakukan dengan menggunakan metode Naive Bayes. Naive Bayes bekerja dengan cara menghitung peluang suatu kalimat terhadap kelas dengan bantuan data latih. Hasil kualitas ringkasan diuji dengan melakukan pencarian nilai precision, recall, F-measure, dan relative utility. Hasil rata-rata F-measure dan relative utility yang didapatkan adalah 0,206538 dan 0,116657.

Salah satu penelitian yang menjadi acuan penelitian yang dilakukan penulis adalah pencarian ringkasan secara ekstraktif oleh Mussina, Aubakirov & Trigo (2018). Ringkasan yang dihasilkan tidak melakukan perubahan terhadap struktur kalimat. Cara yang digunakan untuk mendapatkan ringkasan ekstraktif adalah menggunakan data *statistical*, yaitu dengan menghitung kesamaan *text units*. *Text units* tersebut dapat berupa kata, kalimat atau paragraf. Penelitian tersebut menggunakan dokumen mengenai bencana alam berbahasa Rusia dan Kazakh, yang mana memiliki struktur kalimat yang jelas.

Penelitian Mussina, Aubakirov & Trigo (2018) merepresentasikan dokumen menggunakan metode *TextRank*, dengan kalimat sebagai node dan nilai *similarity* antara kalimat sebagai *edges*. Panjang ringkasan yang diambil adalah sebanyak 30% dari total panjang dokumen. Nilai *similarity* yang digunakan adalah *content overlap*, BM25, dan *substring* terpanjang yang muncul antara dua kalimat. Namun, penelitian tersebut tidak menggunakan metode *PageRank* dalam proses pemeringkatan kalimat, melainkan dengan menjumlahkan seluruh nilai *similarity* yang dimiliki kalimat dan mengurutkannya berdasarkan jumlah nilai terbesar. Secara umum, setiap kalimat saling berkaitan antara satu dengan yang lainnya, namun terkadang ada kalimat yang tidak memiliki kesamaan dengan kalimat lain. Kalimat tersebut tidak memiliki nilai *similarity* sehingga tidak dijadikan ringkasan.

Sebelum melakukan pengambilan ringkasan berdasarkan nilai fungsi *similarity*, penelitian Mussina, Aubakirov & Trigo (2018) mereduksi kalimat berdasarkan nilai *threshold*. Kalimat yang memiliki jumlah nilai *similarity* kurang dari nilai *threshold* tidak dijadikan ringkasan. Pengambilan ringkasan diambil sebanyak 30% dari panjang dokumen berdasarkan kalimat yang memiliki nilai *similarity* terbesar. Kalimat terpilih dijadikan ringkasan berdasarkan urutan sesuai dokumen awal. Evaluasi pada penelitian tersebut dilakukan dengan mencari nilai distribusi *keywords*, dengan nilai rata-rata untuk nilai *similarity* menggunakan *content overlap*, *LongestCommonSubstring*, dan BM25 secara berurutan adalah 0,180; 0,175; dan 0,169.

Penelitian lain yang menjadi acuan penelitian ini adalah percobaan untuk membandingkan hasil ringkasan ekstraktif menggunakan beberapa nilai *similarity* dalam *TextRank* (Barrios et al., 2016). Cara mengidentifikasi hubungan antara kalimat satu dengan kalimat lainnya adalah dengan melakukan perhitungan kata yang sama, *cosine distance*, dan kesamaan *query* yang dianggap penting. Pada penelitian tersebut dokumen direpresentasikan sebagai graf dengan kalimat sebagai *node* dan nilai *similarity* sebagai *edges* (hubungan antara *nodes*). Fungsi *similarity* yang digunakan adalah *Longest Common Substring*, *cosine distance*, BM25, dan BM25+ dengan proses pemeringkatan *PageRank*.

Dokumen yang digunakan pada penelitian Barrios et al., (2016) adalah *Document Understanding Conference (DUC)* yang berjumlah sebanyak 567 dokumen dengan peringkasan sebanyak 20% dari tiap panjang dokumen. Hasil ringkasan dari percobaan tersebut dievaluasi menggunakan metode *ROUGE-N* dengan nilai terbaik didapatkan pada peringkasan menggunakan BM25 dan BM25+.

2.2 Dasar Teori Peringkasan Teks

Ketika jumlah informasi *online* semakin banyak, maka kebutuhan atas sistem yang dapat merangkum satu atau lebih dokumen secara otomatis sangat diperlukan (Radev et al., 2002). Selain dapat membantu dalam mengatasi informasi yang berlebihan, peringkasan otomatis juga berguna dalam penyajian informasi singkat mengingat ukuran perangkat *handy* yang digunakan pembaca (Sankarasubramaniam et al., 2014). Ukuran ringkasan biasanya tidak lebih dari setengah dokumen aslinya. Berdasarkan kebutuhannya ringkasan dokumen terbagi atas (Munot and S. Govilkar, 2013):

1. Metode

a. Abstraktif

Ringkasan yang terdiri dari kalimat baru yang dapat merepresentasikan konteks tulisan dalam bentuk kalimat lain.

b. Ekstraktif

Ringkasan yang terdiri atas kumpulan dari bagian-bagian penting suatu tulisan yang dapat mewakili keseluruhan teks.

2. Konten

a. *Generic*

Ringkasan umum yang tidak bergantung pada syarat apapun.

b. *Query based*

Ringkasan hanya didapatkan berdasarkan *query* yang diinginkan pengguna.

3. Jumlah dokumen

a. *Single document*

Ringkasan yang didapatkan dari satu dokumen.

b. *Multi document*

Ringkasan yang didapatkan dari beberapa dokumen.

4. Bahasa

a. *Monolingual*

Ringkasan yang didapatkan dari dokumen dengan bahasa yang sama.

b. *Multilingual*

Ringkasan yang didapatkan dari beberapa dokumen dengan bahasa yang berbeda.

Dalam penelitian ini peringkasan dilakukan secara ekstraktif dengan memilih kalimat yang dianggap penting dan dapat dijadikan ringkasan. Dokumen yang digunakan merupakan *single document* dengan berbahasa Indonesia.

2.3 Web Scrapping

Untuk mendapatkan dokumen berita pada portal online, sistem yang dibuat dibantu dengan menggunakan metode *web scrapping*. *Web scrapping* mengakses halaman web yang diinginkan dan mengambil data tertentu pada halaman tersebut, lalu mengubahnya sehingga menjadi kumpulan data terstruktur yang dapat digunakan untuk proses komputasi selanjutnya (Boeing and Waddell, 2017). Data yang digunakan dalam melakukan *web scrapping* adalah URL dari halaman portal berita online. *Web scrapping* menerima HTML data dari halaman web tersebut. Data direduksi sehingga hanya menghasilkan judul berita dan konten berita yang digunakan dalam proses peringkasan.

2.4 Preprocessing

Preprocessing dilakukan sebelum proses pembentukan ringkasan dan mempermudah proses peringkasan. Dalam *preprocessing* terdapat tiga tahapan yaitu segmentasi, *cleaning*, tokenisasi, dan *stemming*.

2.4.1 Segmentasi

Segmentasi merupakan proses pemecahan teks dokumen menjadi kalimat-kalimat untuk mempermudah pemrosesan dokumen menjadi potongan-potongan yang lebih kecil. Pada penelitian yang dilakukan, judul pada artikel berita tidak diikutsertakan dalam percobaan. Contoh hasil segmentasi dokumen ditunjukkan pada Tabel 2.1.

Tabel 2.1 Contoh Hasil Segmentasi Dokumen

Index	Kalimat
1	Dia mengambilnya dengan senang hati.
2	Esoknya setelah pulang sekolah, ia melarikan diri.
3	Ibu yang bekerja sebagai pembantu.

2.4.2 Cleaning

Pada proses *cleaning* dilakukan penghilangan tanda baca dan URL karna dianggap tidak mempengaruhi isi dokumen.

2.4.3 Stemming

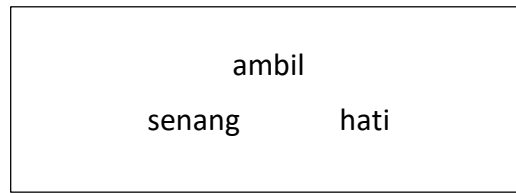
Stemming merupakan metode pembentukan kata atau term menjadi kata dasar. Proses *stemming* dilakukan dengan membuang imbuhan yang terdapat pada term. Contoh hasil stemming dokumen ditunjukkan pada Tabel 2.2.

Tabel 2.2 Contoh Hasil Stemming Dokumen

Index	Kalimat
1	dia ambil dengan senang hati
2	esok telah pulang sekolah ia lari diri
3	ibu yang kerja bagai bantu

2.4.4 Tokenisasi

Tokenisasi adalah metode pemecahan teks kalimat menjadi token-token (*term*) yang berurutan. Pada proses ini juga dilakukan penghilangan angka dan kata hubung atau kata yang dianggap tidak mempengaruhi isi dari konten dokumen menggunakan metode *stopwords removal*. Hal ini juga dilakukan guna meningkatkan performa sistem agar sistem bisa secara efektif dimanfaatkan untuk pengolahan konten yang benar-benar dianggap penting saja. Contoh hasil tokenisasi dokumen ditunjukkan pada Gambar 2.1.



Gambar 2.1 Contoh Hasil Tokenisasi Kalimat Index 1

2.5 Term Weighting

2.5.1 Term Frequency (tf)

Perhitungan *term frequency* merupakan langkah awal dari *term weighting*. *Term frequency* adalah jumlah kemunculan setiap *term* dalam satu kalimat. Contoh hasil perhitungan term frequency ditunjukkan pada Tabel 2.3.

Tabel 2.3 Contoh Hasil Perhitungan Tf

<i>Term</i>	Kalimat 1	Kalimat 2	Kalimat 3
ambil	1	0	0
bantu	0	0	1
esok	0	1	0
hati	1	0	0
kerja	0	0	1
lari	0	1	0
pulang	0	1	0
sekolah	0	1	0
senang	1	0	0

2.5.2 Document Frequency (df)

Document frequency merupakan jumlah dokumen yang memiliki kemunculan sebuah *term*. Nilai *document frequency* digunakan untuk mendapatkan nilai *Inverse document frequency*.

2.5.3 Inverse Document Frequency (idf)

Nilai *inverse document frequency* digunakan untuk perhitungan BM25. Mengacu pada penelitian yang dilakukan (Lv and Zhai, 2011) perhitungan *Idf* ditunjukkan pada Persamaan 2.1 :

$$Idf_t = \log_{10} \frac{N + 1}{dft} \quad (2.1)$$

Keterangan:

N = Total jumlah kalimat
 dft = *Document frequency* tiap *term* dalam satu dokumen

Contoh perhitungan df dan idf ditunjukkan pada Tabel 2.4.

Tabel 2.4 Contoh Hasil Perhitungan Df dan Idf

<i>Term</i>	<i>df</i>	<i>idf</i>
ambil	1	0,60206
bantu	1	0,60206
esok	1	0,60206
hati	1	0,60206
kerja	1	0,60206
lari	1	0,60206
pulang	1	0,60206
sekolah	1	0,60206
senang	1	0,60206

2.6 *TextRank*

TextRank merupakan sebuah algoritme berbasis graf yang digunakan untuk menentukan *node* mana yang paling penting dalam suatu graf (Tarau, 1973). Struktur *TextRank* ditunjukkan pada Gambar 2.2. Penggunaan *TextRank* dapat dilakukan dalam melakukan penarikan keputusan. Proses melakukan *TextRank* dalam peringkasan teks ekstraktif dimulai dengan mengidentifikasi *single document* yang digunakan. Lalu setiap kalimat direpresentasikan sebagai *node* dan hubungan antara kalimat merupakan fungsi *similarity* yang direpresentasikan sebagai *edges*. *Edges* pada graf dapat memiliki arah maupun tidak. Setelah graf terbentuk, maka selanjutnya dilakukan pemeringkatan graf dan mengurutkan *node* yang memiliki nilai pemeringkatan paling tinggi (paling penting). Setelah diurutkan, maka peringkasan dapat diambil berdasarkan peringkat dari kalimat-kalimat tersebut.

2.7 Fungsi *Similarity* BM25

Dalam merepresentasikan dokumen sebagai graf, *edges* didapatkan dari hasil fungsi *similarity* antar kalimat. Fungsi *similarity* didapatkan dari kemiripan isi kalimat satu dengan kalimat lainnya. Kalimat yang merepresentasikan suatu konteks dalam dokumen akan merekomendasikan kalimat lain yang memiliki

konteks yang sama (Tarau, 1973). Fungsi *similarity* yang digunakan dalam penelitian ini adalah BM25. Nilai BM25 didapatkan dari perhitungan bobot *tf* dan *idf* pada setiap kata (*term*). Selain itu juga ditambahkan parameter bebas k_1 dan b dengan nilai k_1 sebesar 1.2 dan b sebesar 0.75 (Manning, C.C.; Ragghavan, P.; Schütze, 2009). Persamaan BM25 dijabarkan pada Persamaan 2.2:

$$RSV_d = \sum_{t \in q} Idf \cdot \frac{(k_1 + 1)tf_{td}}{k_1 \left((1 - b) + b \left(\frac{L_d}{L_{ave}} \right) \right) + tf_{td}} \quad (2.2)$$

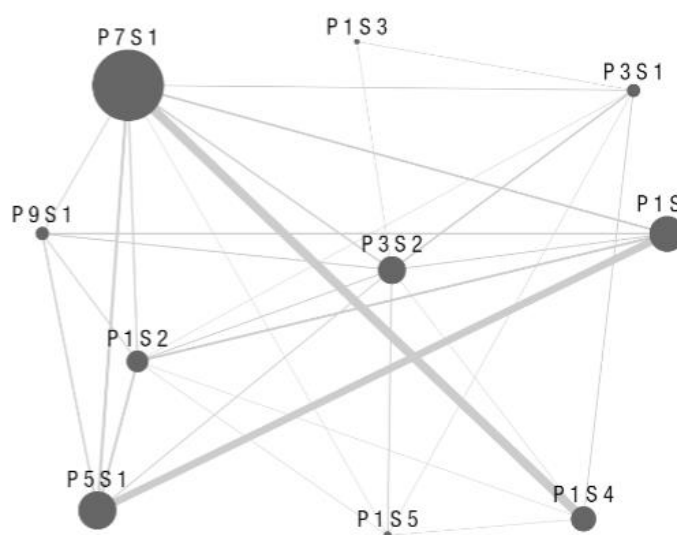
Keterangan:

$\sum_{t \in q} Idf$ = Nilai *idf term t*

k_1 dan b = Parameter penskalaan terhadap *tf* dan panjang dokumen

tf_{td} = Frekuensi term *t* pada kalimat *d*

L_d dan L_{ave} = Panjang kalimat *d* dan rata-rata dari panjang seluruh koleksi kalimat



Gambar 2.2 Struktur TextRank Sebagai Graf

Sumber: (Demo Page for TextRank Algorithm, 2020)

2.8 PageRank

PageRank adalah metode yang digunakan dalam pemeringkatan graf. *PageRank* digunakan oleh Google untuk menentukan tingkat kepentingan halaman web. *PageRank* merupakan nilai numerical yang menyatakan seberapa penting sebuah halaman web di internet. Singkatnya, perhitungan nilai tersebut bertambah bila halaman tersebut muncul sebagai sebuah *hyperlink* di sebuah halaman web lainnya. Semakin besar nilai yang dimiliki, maka semakin penting web tersebut. Begitu juga dengan kalimat yang saling berhubungan satu sama lain dalam sebuah graf. Kalimat yang penting akan memiliki nilai *PageRank* yang besar.

Inisialisasi awal nilai *PageRank* tiap kalimat ditentukan secara random mulai dari 0 hingga 1 (Brin and Page, 1998). Lalu sejumlah iterasi dilakukan untuk melakukan *update* bobot *PageRank* di tiap kalimat. Persamaan *PageRank* dijabarkan pada Persamaan 2.3:

$$WS(V_i) = (1 - d) + d * \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} WS(V_j) \quad (2.3)$$

Keterangan:

$WS(V_i)$ = Bobot nilai *PageRank* kalimat *i*

d = *Dampening factor*

$V_j \in In(V_i)$ = Kalimat *j* yang berhubungan dengan kalimat *i*

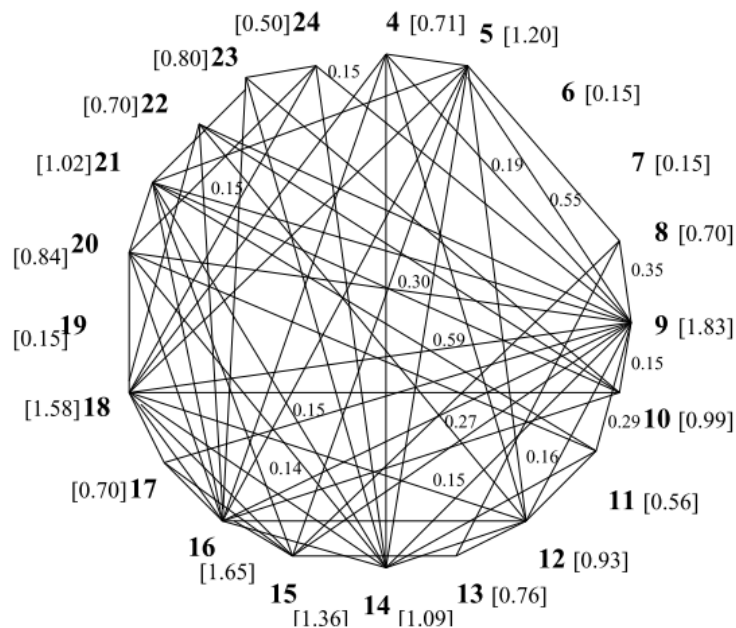
w_{ji} = Nilai fungsi *similarity* antara kalimat *j* dan *i*

$V_k \in Out(V_j)$ = Kalimat *k* yang berhubungan dengan kalimat *j*

w_{jk} = Nilai fungsi *similarity* antara kalimat *j* dan *k*

$WS(V_j)$ = Bobot nilai *PageRank* kalimat *j*

Dampening factor (d) adalah nilai yang telah dihitung oleh Google Engineers dalam sistem *PageRank* untuk memastikan bahwa bobot *node* akan konvergen pada satu nilai. Nilai *dampening factor* bisa didapatkan dari angka random mulai dari nol hingga satu, namun 0.85 telah menjadi nilai yang umum saat menetapkan nilai *dampening factor* (Brin and Page, 1998). Pada akhir perhitungan graf kalimat dapat diilustrasikan seperti pada Gambar 2.3 .



Gambar 2.3 Struktur *TextRank* Setelah Perhitungan *Pagerank*

Sumber: (Tarau, 1973)

2.9 Compression Rate

Compression rate digunakan sebagai acuan panjang ringkasan yang dihasilkan sistem peringkasan otomatis. Semakin kecil kompresi rasio (*compression rate*) yang digunakan maka semakin banyak informasi yang dibuang, namun semakin besar kompresi rasio yang digunakan maka semakin tidak penting informasi yang dihasilkan (Alguliev & Aliguliyev, 2009). Besar kompresi rasio yang diterima untuk peringkasan adalah sebesar 5-30% (Hahn & Mani, 2000).

2.10 Evaluasi

Pengujian sistem peringkasan ini dilakukan dengan membandingkan nilai *precision*, *recall*, dan *f-measure* untuk setiap hasil ringkasan (Steinberger and Ježek, 2009). Proses perhitungan tersebut dibantu dengan menggunakan *confussion matrix*.

a. Confussion Matrix

Confusion Matrix merupakan informasi mengenai tingkat relevansi hasil peringkasan yang didapatkan oleh sistem. Contoh *confusion matrix* ditunjukkan pada Tabel 2.5.

Tabel 2.5 Contoh Confusion Matrix

	<i>Relevant</i>	<i>Not relevant</i>
<i>Retrieved</i>	A	B
<i>Not retrieved</i>	C	D

Keterangan :

A = Semua data yang dihasilkan sistem dan relevan bagi *expert*

B = Semua data yang dihasilkan sistem dan tidak relevan bagi *expert*

C = Semua data yang relevan bagi *expert* dan tidak dihasilkan sistem

D = Semua data yang tidak relevan bagi *expert* dan tidak dihasilkan sistem

b. Precision

Merupakan nilai ketepatan antara informasi yang dihasilkan sistem dengan hasil informasi yang seharusnya (dianggap benar). Persamaan *precision* ditunjukan pada Persamaan 2.5:

$$Precision = \frac{A}{A + B} \quad (2.5)$$

Keterangan:

A = Semua data yang dihasilkan sistem dan relevan bagi *expert*

B = Semua data yang dihasilkan sistem dan tidak relevan bagi *expert*

c. Recall

Merupakan tingkat keberhasilan sistem dalam peringkasan otomatis yang menentukan berapa proporsi kalimat yang dipilih oleh pakar yang juga dipilih oleh sistem. Persamaan *recall* ditunjukkan pada Persamaan 2.6:

$$Recall = \frac{A}{A + C} \quad (2.6)$$

Keterangan:

A = Semua data yang dihasilkan sistem dan relevan bagi *expert*

C = Semua data yang relevan bagi *expert* dan tidak dihasilkan sistem

d. F-Measure

Merupakan pengukuran yang menilai timbal balik antara *precision* dan *recall*. Persamaan *F-Measure* ditunjukkan pada Persamaan 2.7:

$$F - measure = \frac{2 \times P \times R}{P + R} \quad (2.7)$$

Keterangan:

P = Precision

R = Recall

BAB 3 METODOLOGI

3.1 Jenis Penelitian

Penelitian yang diajukan bersifat Non-implementatif analitik. Hasil dari penelitian ini akan dibandingkan dengan ringkasan yang diambil oleh pakar untuk mengetahui kualitas ringkasan. Hasil tersebut kemudian dianalisis menggunakan metode evaluasi.

3.2 Pengambilan Data

Jumlah populasi pada penelitian ini termasuk ke dalam kategori tak terbatas karena dokumen yang digunakan merupakan artikel berita online berbahasa Indonesia. Populasi tak terbatas yaitu sumber datanya tidak dapat ditentukan batasan-batasannya sehingga relatif tidak dapat ditentukan dalam bentuk angka. Teknik pengambilan sampel yang digunakan adalah *simple random sampling*. *simple random sampling* merupakan pengambilan sampel dari anggota populasi secara acak tanpa memperhatikan tingkatan, karena populasi bersifat homogen (sama).

Instrumen pengumpulan dokumen pada penelitian ini adalah peneliti sendiri, yang mana dokumen diambil dari artikel berita online BBC Indonesia yang dapat diakses melalui situs resmi BBC Indonesia. Jumlah dokumen yang diambil sebagai dokumen pengujian sistem adalah sebanyak 10 artikel berita mengenai bencana alam dan lingkungan. Jumlah artikel berita yang cukup sedikit diambil karena penelitian ini bersifat mendalam dengan membandingkan hasil ringkasan menggunakan beberapa jenis *compression rate*.

3.3 Metode Penelitian

Penelitian ini terbagi atas tiga tahap, yaitu *web scrapping*, pembentukan ringkasan dan evaluasi ringkasan yang dihasilkan. Tahap *web scrapping* dilakukan untuk mengambil dokumen berita pada halaman *website* BBC Indonesia dengan menggunakan URL. Proses *web scrapping* dilakukan demi meningkatkan efisiensi sistem peringkasan otomatis yang dibuat. Tahap pembentukan ringkasan terdiri atas tahapan *preprocessing*, *term weighting*, penerapan *TextRank* dan *PageRank*, dan pembentukan ringkasan. Hasil ringkasan yang didapatkan dari penelitian ini adalah kalimat yang memiliki nilai *PageRank* yang besar setelah diurutkan. Kalimat tersebut dianggap penting dan dapat merepresentasikan keseluruhan dokumen sehingga dapat dijadikan ringkasan. Mengacu pada penelitian yang dilakukan (Yeh et al., 2005) total kalimat yang diambil sebagai ringkasan ialah sebanyak 5%-30% dari total keseluruhan kalimat. Untuk tahap evaluasi, hasil ringkasan yang telah didapatkan sistem dibandingkan dengan hasil ringkasan manual yang dilakukan oleh pakar. Secara umum, diagram proses peringkasan teks pada penelitian ini ditunjukkan pada Gambar 3.1.

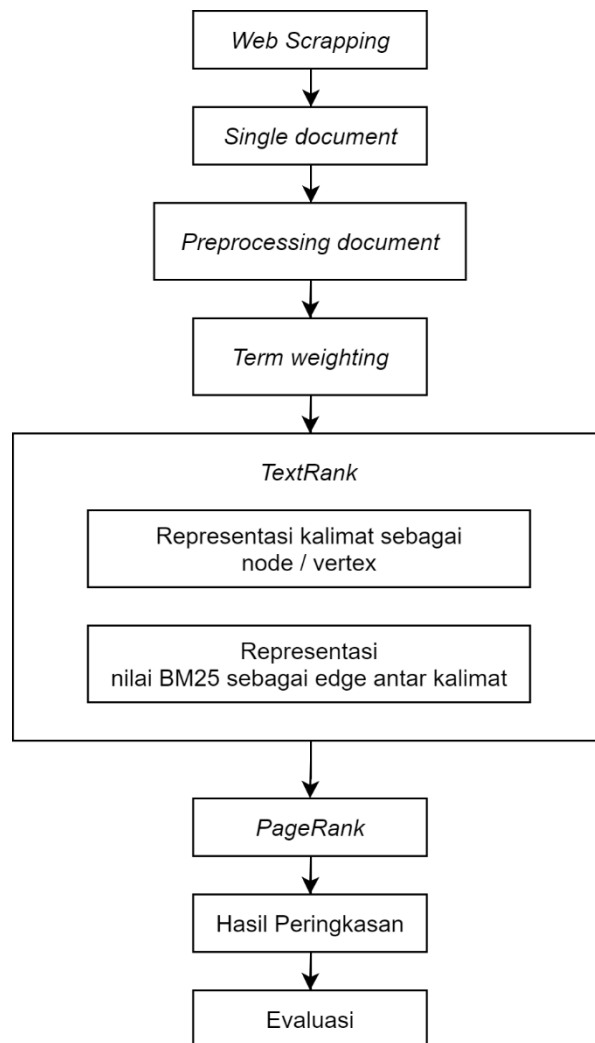
3.4 Peralatan Pendukung

Lingkungan implementasi system *hardware* meliputi:

1. Processor 2.00 GHz
2. RAM 8.00 GB

Spesifikasi *software* meliputi:

1. Operating System Windows 10 64-bit
2. Jupyter Notebook



Gambar 3.1 Gambaran Umum Proses Peringkasan *Single Document*

3.5 Teknik Analisis Data

Pengujian hasil ringkasan yang dihasilkan sistem dilakukan oleh dosen Bahasa Indonesia dari Universitas Brawijaya. Dosen Bahasa Indonesia tersebut dijadikan pakar dan diminta untuk memilih kalimat yang dianggap penting dan dapat dijadikan ringkasan pada sekumpulan dokumen yang diujikan. Hasil pemilihan

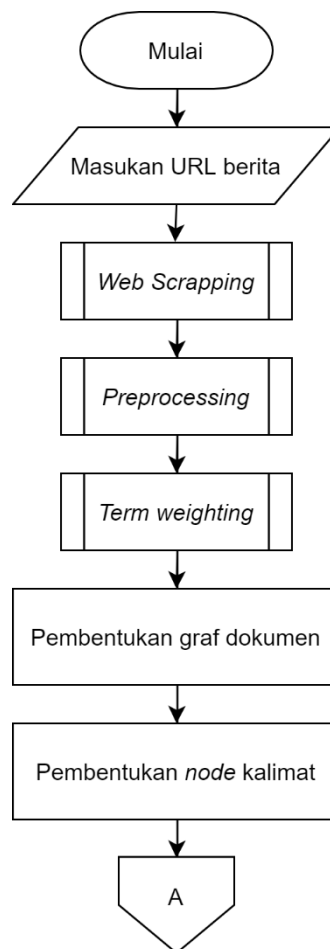
kalimat pada tiap dokumen akan dibandingkan dengan hasil peringkasan otomatis yang didapatkan oleh sistem yang dibuat menggunakan *TextRank*. Perhitungan *precision*, *recall*, dan *F-Measure* digunakan untuk mengetahui kualitas ringkasan otomatis tersebut.

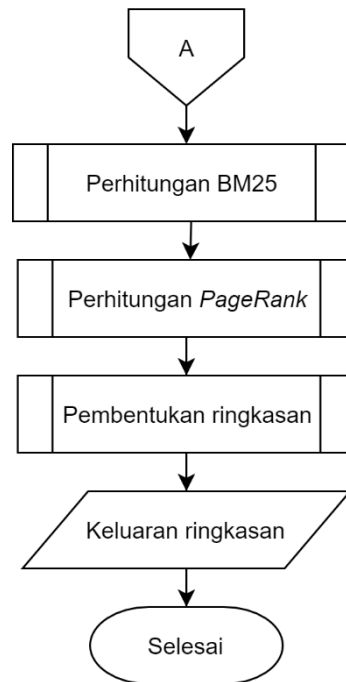
BAB 4 PERANCANGAN

4.1 Perancangan Sistem

4.1.1 Perancangan Umum Sistem Peringkasan Otomatis

Secara umum perancangan sistem dimulai dengan masukan URL halaman berita yang diringkaskan. Proses *web scrapping* dilakukan untuk mendapatkan konten dokumen berita dari URL tersebut. Konten berita digunakan dalam *preprocessing* agar dokumen terpecah menjadi potongan kalimat yang hanya berisikan *term-term* penting. Hasil *term* digunakan untuk menghitung nilai bobot sesuai metode *tf* dan *idf*. Setelah mendapatkan bobot tiap *term*, dokumen direpresentasikan sebagai graf dan kalimat direpresentasikan sebagai *node* dalam graf sesuai dengan metode *TextRank*. Perhitungan BM25 dilakukan untuk menghitung nilai kemiripan antara kalimat dan pemeringkatan kalimat dilakukan menggunakan metode *PageRank*. Selanjutnya ringkasan dipilih berdasarkan pemeringkatan tersebut. Diagram alir sistem ditunjukkan pada Gambar 4.1.





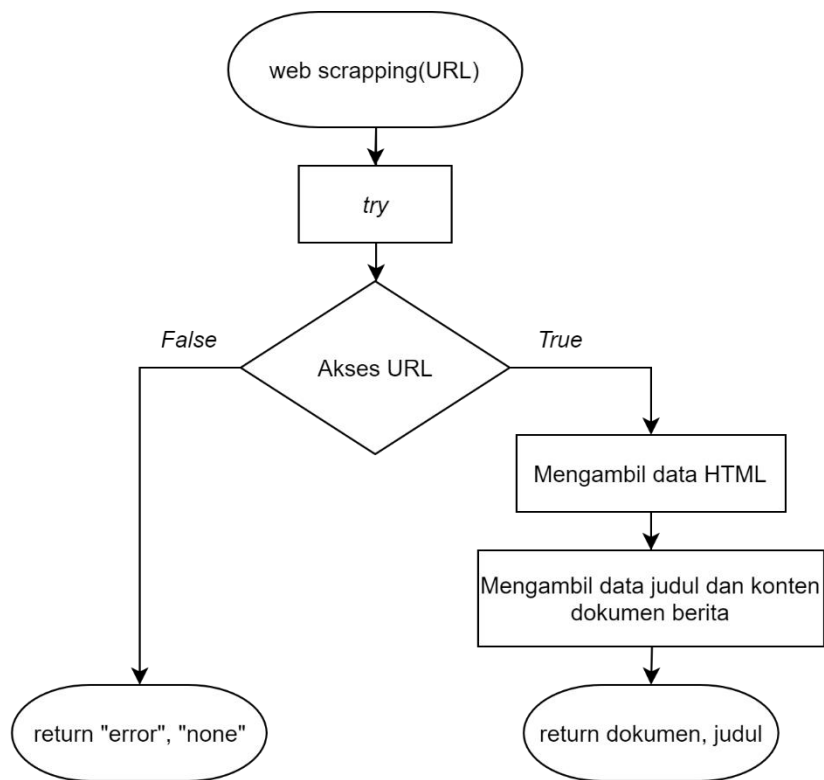
Gambar 4.1 Diagram Alir Sistem

4.1.2 Web Scrapping

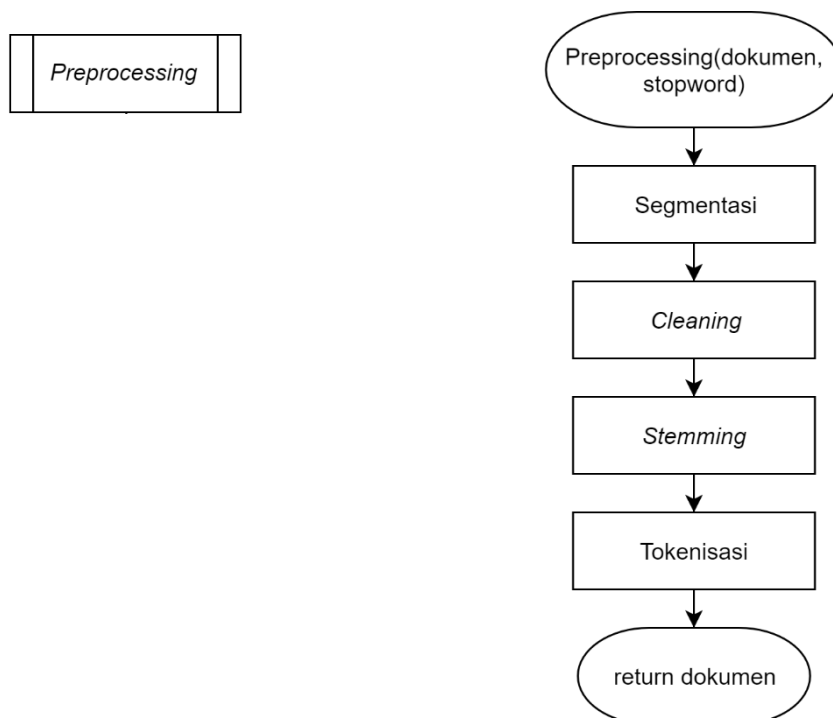
Proses *web scrapping* dilakukan untuk meningkatkan performa dan efisiensi sistem peringkasan otomatis, sehingga sistem dapat digunakan dengan memasukkan URL dari halaman portal berita *online*. *Web scrapping* bekerja dengan mengakses URL yang diberikan pengguna dan mengambil data HTML pada halaman web tersebut. Jika akses URL yang dilakukan tidak berhasil maka sistem mengembalikan pesan error. Selanjutnya dilakukan seleksi data untuk mengambil bagian judul dan konten dengan memilih *tag div*. Konten dokumen berita yang berhasil diambil digunakan untuk proses peringkasan. Diagram alir *web scrapping* ditunjukkan pada Gambar 4.2 .

4.1.3 Perancangan Preprocessing

Pada proses *preprocessing*, dokumen berita dibersihkan dengan cara menghilangkan URL, angka, symbol, dan *stopword*. *Stopword* merupakan kumpulan kata yang dianggap tidak penting dan tidak mempengaruhi isi dokumen. Setelah pembersihan dokumen, tiap kalimat dipecah menjadi *term* atau kata. Alur *preprocessing* ditunjukkan pada Gambar 4.3.



Gambar 4.2 Diagram Alir *Web Scrapping*



Gambar 4.3 Diagram Alir *Preprocessing*

4.1.3.1 Segmentasi

Segmentasi merupakan proses pemisahan dokumen menjadi potongan-potongan kalimat. Proses segmentasi dilakukan dengan bantuan *library* NLTK dengan modul *sent_tokenize*. Hal ini dilakukan agar pemisahan kalimat tidak hanya mengandalkan tanda titik karna pemisahan bisa terjadi ditengah nama orang atau jalan memiliki titik karna adanya singkatan atau gelar.

4.1.3.2 Cleaning

Pada proses ini dilakukan pembersihan dokumen dengan menghilangkan URL dan simbol. Hal ini diperlukan untuk menghilangkan bagian kalimat yang dianggap tidak penting.

4.1.3.3 Stemming

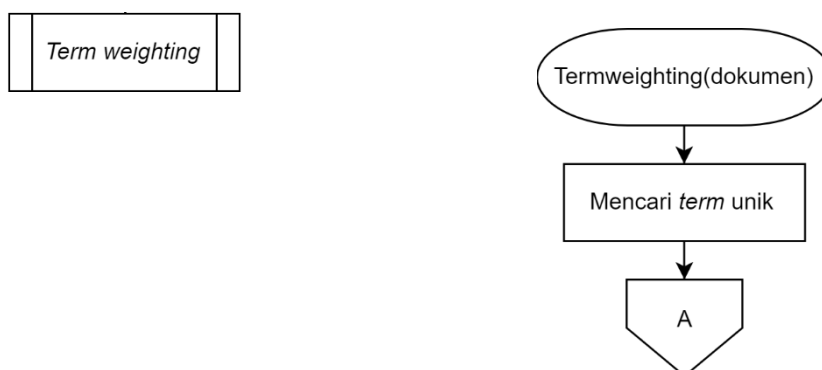
Stemming dilakukan untuk menghilangkan imbuhan pada kata dalam kalimat sehingga menghasilkan kata dasar. Proses stemming dilakukan dengan bantuan *library* stemming bahasa Indonesia Stemmer Sastrawi. Hasil *stemming* digunakan untuk proses berikutnya yaitu tokenisasi.

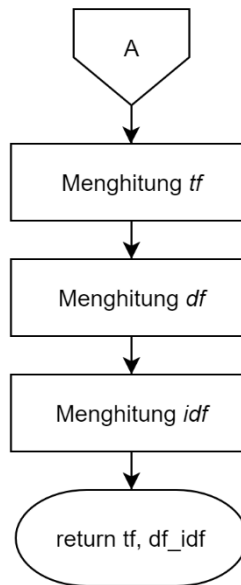
4.1.3.4 Tokenisasi

Proses tokenisasi dilakukan untuk memisahkan kalimat menjadi potongan-potongan *term* atau kata. Pada proses ini juga dilakukan penghilangan angka dan *stopword removal*, yaitu penghilangan kata yang dianggap tidak penting dan tidak mempengaruhi kata penting yang berkaitan dengan konteks dokumen. Jenis kata yang biasanya dihilangkan pada proses *stopword removal* adalah kata hubung.

4.1.4 Perancangan *Term Weighting*

Proses *term weighting* dilakukan untuk mendapatkan bobot *Inverse Document Frequency (Idf)* yang digunakan pada perhitungan fungsi *similarity* BM25. *Term weighting* dimulai dengan mencari *term* unik pada dokumen, lalu menghitung jumlah munculnya *term*/kata dalam satu kalimat yang disebut juga dengan *frequency (tf)*. Nilai *tf* digunakan untuk menghitung *document frequency (df)* yaitu jumlah dokumen yang terdapat kemunculan sebuah *term*, dan mencari nilai *Inverse Document Frequency (Idf)* menggunakan Persamaan 2.1. Alur *term weighting* ditunjukkan pada Gambar 4.4.

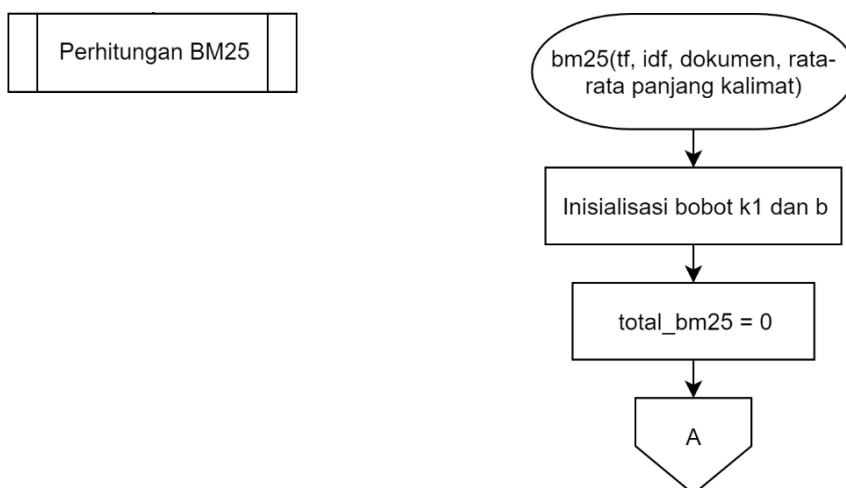




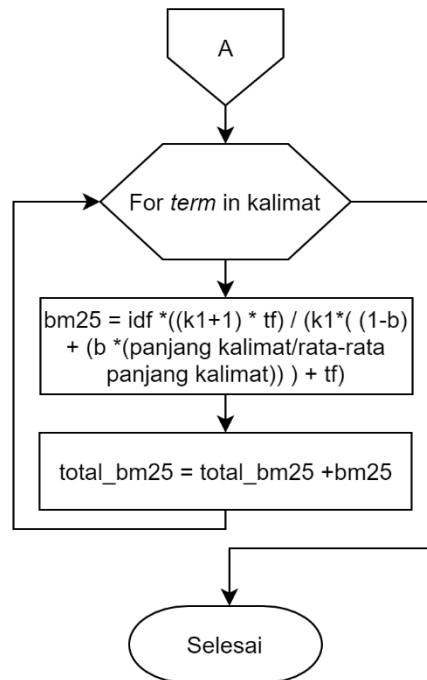
Gambar 4.4 Diagram Alir *Term Weighting*

4.1.5 Perancangan Perhitungan BM25

Perhitungan nilai BM25 dilakukan dengan menetapkan nilai parameter k_1 dan b seperti pada Persamaan 2.2. Untuk setiap kalimat dilanjutkan dengan perhitungan BM25 dengan kalimat lain dalam satu dokumen. Kalimat pertama dijadikan *query* dalam perhitungan BM25 lalu sistem menghitung kemiripan kalimat tersebut dengan sisa kalimat lainnya. Hal tersebut terus dilakukan hingga mencapai kalimat terakhir. Semakin besar nilai BM25 antara satu kalimat dengan kalimat lain, maka semakin mirip topik yang dibicarakan pada kedua kalimat tersebut. Perancangan perhitungan BM25 ditunjukkan pada Gambar 4.5.



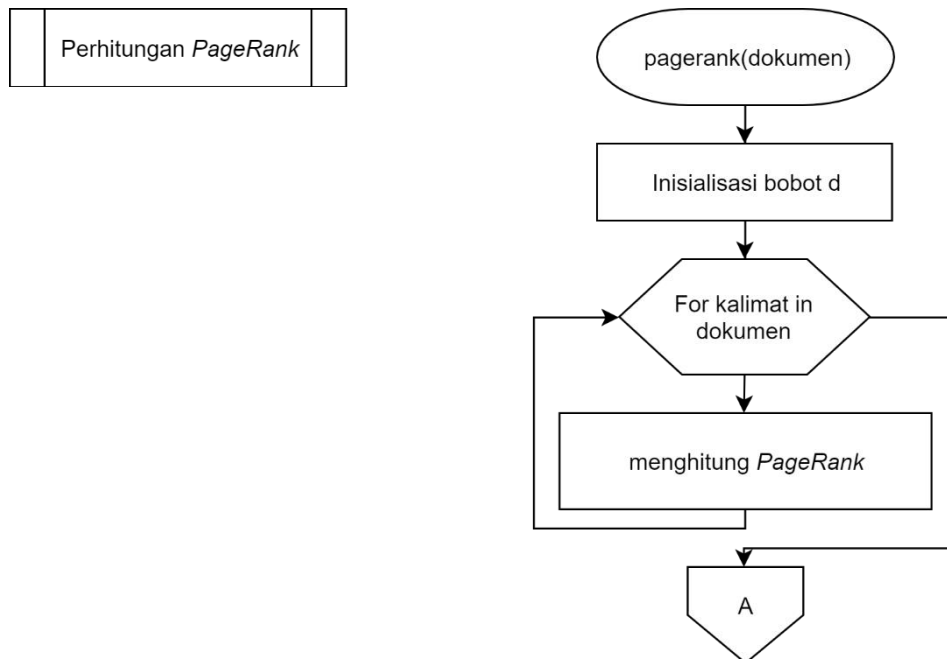
Perhitungan BM25

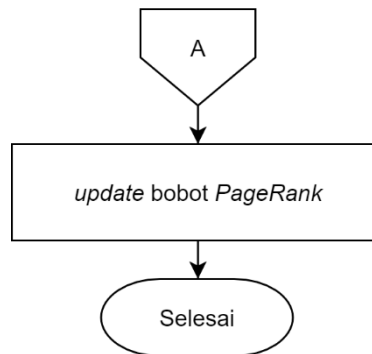


Gambar 4.5 Diagram Alir Perhitungan BM25

4.1.6 Perancangan Perhitungan *PageRank*

PageRank merupakan teknik pemeringkatan berbasis graf dengan memanfaatkan bobot yang dimiliki antara suatu node dengan node yang lain sesuai dengan Persamaan 2.3. Proses perhitungan *PageRank* dilakukan terhadap setiap *node* kalimat sebanyak jumlah iterasi yang ditentukan. Bobot awal *PageRank* untuk setiap kalimat ditetapkan secara random antara nol hingga satu. Perancangan perhitungan nilai *PageRank* ditunjukkan pada Gambar 4.6.

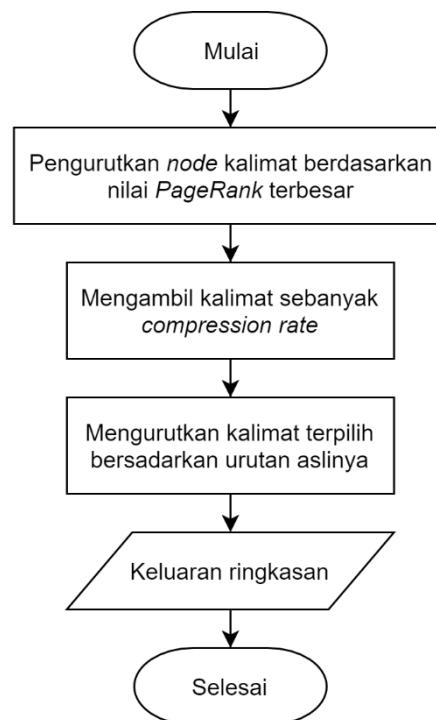
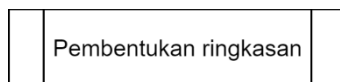




Gambar 4.6 Diagram Alir Perhitungan *PageRank*

4.1.7 Perancangan Mendapatkan Ringkasan

Ringkasan diambil berdasarkan urutan nilai *PageRank* terbesar dari seluruh *node* kalimat. Kalimat dipilih sebanyak *compression rate* dari seluruh panjang dokumen dan diurutkan sesuai dengan urutan asli dokumennya. Perancangan pengambilan ringkasan ditunjukkan pada Gambar 4.7.



Gambar 4.7 Diagram Alir Pembentukan Ringkasan

4.2 Perhitungan Manual

Perhitungan manual dilakukan untuk membuktikan dan mempermudah proses implementasi system yang dibuat.

4.2.1 Data Perhitungan Manual

Proses perhitungan manual ini dilakukan menggunakan satu dokumen dengan total 10 kalimat. Dokumen tersebut ditunjukkan pada Gambar 4.8.

Sejumlah sekolah di Kota Palembang, Sumatera Selatan, memutuskan memulangkan siswa-siswa mereka lantaran kabut asap semakin tebal menyelimuti kota tersebut. pada Senin (14/10) pagi. Kepala SMP Negeri 7 Palembang, Siti Zubaida, mengatakan keputusan pemulangan ditempuh sesuai dengan instruksi Dinas Pendidikan Kota Palembang. "Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan berdampak buruk terhadap siswa, oleh karenanya atas instruksi Kadiknas Kota Palembang melalui pesan WA Grup meminta siswa dipulangkan dan belajar di rumah masing-masing saja," jelas Siti kepada radio Elshinta. Hal ini diamini Kepala Dinas Pendidikan Kota Palembang, Ahmad Zulinto, yang menyampaikan surat edaran ke semua sekolah. "Hari ini seluruh TK hingga SMP negeri dan swasta sederajat diliburkan, untuk besok dan seterusnya akan diberikan edaran lebih lanjut," kata Ahmad Zulinto kepada kantor berita Antara. Menurutnya, kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan, kegiatan belajar mengajar di tingkat TK sampai SMP akan tetap diliburkan. akan tetapi, sebagaimana dipaparkan Kepala Dinas Pendidikan Sumatera Selatan, Widodo, kegiatan belajar mengajar di daerah yang tidak terdampak kabut asap tetap berlangsung. "Daerah yang tidak terdampak kabut asap tetap normal tetap belajar, untuk daerah yang terkategori sedang tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker, bagi daerah terkategori parah maka siswa diberikan tugas dengan memaksimalkan kelas daring," kata Widodo kepada Antara. Hal itu belakangan dibenarkan Agus Wibowo, selaku Kepala Pusat Data, Informasi dan Humas Badan Nasional Penanggulangan Bencana (BNPB). "Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud, TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers.

Gambar 4.8 Dokumen Perhitungan Manualisasi

4.2.2 Preprocessing

4.2.2.1 Segmentasi

Proses pemotongan dokumen menjadi kalimat-kalimat. Hasil segmentasi ditunjukkan pada Tabel 4.1.

Tabel 4.1 Hasil Segmentasi

Index	Kalimat
1	Sejumlah sekolah di Kota Palembang, Sumatera Selatan, memutuskan memulangkan siswa-siswa mereka lantaran kabut asap semakin tebal menyelimuti kota tersebut pada Senin (14/10) pagi.

2	Kepala SMP Negeri 7 Palembang, Siti Zubaida, mengatakan keputusan pemulangan ditempuh sesuai dengan instruksi Dinas Pendidikan Kota Palembang.
3	"Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan berdampak buruk terhadap siswa, oleh karenanya atas instruksi Kadiknas Kota Palembang melalui pesan WA Grup meminta siswa dipulangkan dan belajar di rumah masing-masing saja," jelas Siti kepada radio Elshinta.
4	Hal ini diamini Kepala Dinas Pendidikan Kota Palembang, Ahmad Zulinto, yang menyampaikan surat edaran ke semua sekolah.
5	"Hari ini seluruh TK hingga SMP negeri dan swasta sederajat diliburkan, untuk besok dan seterusnya akan diberikan edaran lebih lanjut," kata Ahmad Zulinto kepada kantor berita Antara.
6	Menurutnya, kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan, kegiatan belajar mengajar di tingkat TK sampai SMP akan tetap diliburkan.
7	Akan tetapi, sebagaimana dipaparkan Kepala Dinas Pendidikan Sumatera Selatan, Widodo, kegiatan belajar mengajar di daerah yang tidak terdampak kabut asap tetap berlangsung.
8	"Daerah yang tidak terdampak kabut asap tetap normal tetap belajar, untuk daerah yang terkategori sedang tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker, bagi daerah terkategori parah maka siswa diberikan tugas dengan memaksimalkan kelas daring," kata Widodo kepada Antara.
9	Hal itu belakangan dibenarkan Agus Wibowo, selaku Kepala Pusat Data, Informasi dan Humas Badan Nasional Penanggulangan Bencana (BNPB).
10	"Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud, TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers.

4.2.2.2 Cleaning

Proses pembersihan kalimat dari URL dan simbol. Hasil cleaning dapat dilihat pada Tabel 4.2.

Tabel 4.2 Hasil *Cleaning*

Index	Kalimat
1	Sejumlah sekolah di Kota Palembang Sumatera Selatan memutuskan memulangkan siswa siswa mereka lantaran kabut asap semakin tebal menyelimuti kota tersebut pada Senin pagi
2	Kepala SMP Negeri Palembang Siti Zubaida mengatakan keputusan pemulangan ditempuh sesuai dengan instruksi Dinas Pendidikan Kota Palembang
3	Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan berdampak buruk terhadap siswa oleh karenanya atas instruksi Kadiknas Kota Palembang melalui pesan WA Grup meminta siswa dipulangkan dan belajar di rumah masing masing saja jelas Siti kepada radio Elshinta
4	Hal ini diamini Kepala Dinas Pendidikan Kota Palembang Ahmad Zulinto yang menyampaikan surat edaran ke semua sekolah
5	Hari ini seluruh TK hingga SMP negeri dan swasta sederajat diliburkan untuk besok dan seterusnya akan diberikan edaran lebih lanjut kata Ahmad Zulinto kepada kantor berita Antara
6	Menurutnya kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan kegiatan belajar mengajar di tingkat TK sampai SMP akan tetap diliburkan
7	Akan tetapi sebagaimana dipaparkan Kepala Dinas Pendidikan Sumatera Selatan Widodo kegiatan belajar mengajar di daerah yang tidak terdampak kabut asap tetap berlangsung
8	Daerah yang tidak terdampak kabut asap tetap normal tetap belajar untuk daerah yang terkategori sedang tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker bagi daerah terkategori parah maka siswa diberikan tugas dengan memaksimalkan kelas daring kata Widodo kepada Antara
9	Hal itu belakangan dibenarkan Agus Wibowo selaku Kepala Pusat Data Informasi dan Humas Badan Nasional Penanggulangan Bencana BNPB
10	Melalui pesan digital Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud TK SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan sebut Agus dalam siaran pers

4.2.2.3 Stemming

Merupakan proses penghilangan imbuhan yang terdapat dalam kalimat. Hasil stemming dapat dilihat pada Tabel 4.3.

Tabel 4.3 Hasil Stemming

Index	Kalimat
1	jumlah sekolah di kota Palembang Sumatera Selatan putus pulang siswa siswa mereka lantaran kabut asap makin tebal limut kota sebut pada Senin pagi
2	kepala SMP Negeri 7 Palembang Siti Zubaida kata putus pulang tempuh sesuai dengan instruksi Dinas Didik Kota Palembang
3	pagi ini kami pulang siswa karena lihat kabut asap yang tebal dan dampak buruk hadap siswa oleh karena atas instruksi Kadiknas Kota Palembang lalu pesan wa grup minta siswa pulang dan ajar di rumah masing masing saja jelas Siti kepada radio Elshinta
4	hal ini amin kepala Dinas Didik Kota Palembang Ahmad Zulinto yang sampai surat edar ke semua sekolah
5	hari ini seluruh TK hingga SMP Negeri dan swasta derajat libur untuk besok dan terus akan beri edar lebih lanjut kata Ahmad Zulinto kepada kantor berita Antara
6	turut kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan giat ajar ajar di tingkat TK sampai SMP akan tetap libur
7	akan tetapi bagaimana papar kepala Dinas Didik Sumatera Selatan Widodo giat ajar ajar di daerah yang tidak dampak kabut asap tetap langsung
8	daerah yang tidak dampak kabut asap tetap normal tetap ajar untuk daerah yang kategori sedang tetap ajar namun jam masuk sekolah undur dan kami himbau pakai masker bagi daerah kategori parah maka siswa beri tugas dengan maksimal kelas daring kata Widodo kepada Antara
9	hal itu belakang benar Agus Wibowo selaku kepala Pusat Data Informasi dan Humas Badan Nasional Tanggulang Bencana BNPB
10	lalu pesan digital kepala Dinas Didik Kota Palembang instruksi giat ajar ajar di tingkat PAUD TK SD dan SMP Negeri dan swasta libur hingga batas yang belum tentu sebut Agus dalam siaran pers

4.2.2.4 Tokenisasi

Pada proses ini dilakukan pemisahan tiap kata dan menghilangkan kata yang terdapat dalam *stopword* dan angka. Hasil tokenisasi dapat dilihat pada Tabel 4.4.

Tabel 4.4 Hasil Tokenisasi

Index	Kalimat
1	sekolah', 'kota', 'palembang', 'sumatera', 'selatan', 'putus', 'pulang', 'siswa', 'lantar', 'kabut', 'asap', 'tebal', 'limut', 'senin', 'pagi'
2	kepala', 'smp', 'negeri', 'palembang', 'siti', 'zubaida', 'putus', 'pulang', 'tempuh', 'sesuai', 'instruksi', 'dinas', 'didik', 'kota'
3	pagi', 'pulang', 'siswa', 'lihat', 'kabut', 'asap', 'tebal', 'dampak', 'buruk', 'hadap', 'instruksi', 'kadiknas', 'kota', 'palembang', 'pesan', 'wa', 'grup', 'ajar', 'rumah', 'siti', 'radio', 'elshinta'
4	amin', 'kepala', 'dinas', 'didik', 'kota', 'palembang', 'ahmad', 'zulinto', 'surat', 'edar', 'sekolah'
5	tk', 'smp', 'negeri', 'swasta', 'derajat', 'libur', 'besok', 'edar', 'ahmad', 'zulinto', 'kantor', 'berita'
6	kualitas', 'udara', 'palembang', 'buruk', 'giat', 'ajar', 'tingkat', 'tk', 'smp', 'libur'
7	papar', 'kepala', 'dinas', 'didik', 'sumatera', 'selatan', 'widodo', 'giat', 'ajar', 'daerah', 'dampak', 'kabut', 'asap', 'langsung'
8	daerah', 'dampak', 'kabut', 'asap', 'normal', 'ajar', 'kategori', 'jam', 'masuk', 'sekolah', 'undur', 'himbau', 'pakai', 'masker', 'parah', 'siswa', 'tugas', 'maksimal', 'kelas', 'daring', 'widodo'
9	agus', 'wibowo', 'kepala', 'pusat', 'data', 'informasi', 'humas', 'badan', 'nasional', 'tanggulang', 'bencana', 'bnpb'
10	pesan', 'digital', 'kepala', 'dinas', 'didik', 'kota', 'palembang', 'instruksi', 'giat', 'ajar', 'tingkat', 'paud', 'tk', 'sd', 'smp', 'negeri', 'swasta', 'libur', 'batas', 'agus', 'siar', 'pers'

4.2.3 Term Weighting

4.2.3.1 Term Frequency (tf)

Pada proses ini dihitung jumlah kemunculan tiap kata pada setiap kalimat. Hasil perhitungan *tf* dapat dilihat pada Tabel 4.5.

Tabel 4.5 Hasil Perhitungan *tf*

<i>Term</i>	Kalimat									
	1	2	3	4	5	6	7	8	9	10
agus	0	0	0	0	0	0	0	0	1	1
ahmad	0	0	0	1	1	0	0	0	0	0
ajar	0	0	1	0	0	2	2	2	0	2
amin	0	0	0	1	0	0	0	0	0	0
asap	1	0	1	0	0	0	1	1	0	0
badan	0	0	0	0	0	0	0	0	1	0
batas	0	0	0	0	0	0	0	0	0	1
bencana	0	0	0	0	0	0	0	0	1	0
..
..
udara	0	0	0	0	0	1	0	0	0	0
undur	0	0	0	0	0	0	0	1	0	0
wa	0	0	1	0	0	0	0	0	0	0
wibowo	0	0	0	0	0	0	0	0	1	0
widodo	0	0	0	0	0	0	1	1	0	0
zubaida	0	1	0	0	0	0	0	0	0	0
zulinto	0	0	0	1	1	0	0	0	0	0
Panjang kalimat	17	15	25	11	12	11	15	25	12	23
Rata-rata panjang kalimat	16,6									

4.2.3.2 Document Frequency (*df*)

Pada proses ini dihitung jumlah kemunculan kata pada seluruh kalimat. Hasil perhitungan *df* dapat dilihat pada Tabel 4.6.

Tabel 4.6 Hasil Perhitungan *df*

<i>Term</i>	<i>df</i>
agus	2
ahmad	2
ajar	5
amin	1
asap	4
badan	1
batas	1
bencana	1
..	..
..	..
udara	1

undur	1
wa	1
wibowo	1
widodo	2
zubaida	1
zulinto	2

4.2.3.3 Inverse Document Frequency (*idf*)

Perhitungan *idf* dilakukan berdasarkan Persamaan 2.1 dengan memanfaatkan hasil dari perhitungan *df* sebelumnya. Hasil perhitungan *idf* dapat dilihat pada Tabel 4.7.

Contoh perhitungan *idf* untuk kata 'agus', 'ahmad', dan 'ajar':

$$Idf_t = \log_{10} \frac{N + 1}{df_t}$$

$$\begin{aligned} Idf_{agus} &= \log_{10} \frac{N + 1}{df_{agus}} \\ &= \log_{10} \frac{10 + 1}{2} \\ &= 0,740362689 \end{aligned}$$

$$\begin{aligned} Idf_{ahmad} &= \log_{10} \frac{N + 1}{df_{ahmad}} \\ &= \log_{10} \frac{10 + 1}{2} \\ &= 0,740362689 \end{aligned}$$

$$\begin{aligned} Idf_{ajar} &= \log_{10} \frac{N + 1}{df_{ajar}} \\ &= \log_{10} \frac{10 + 1}{5} \\ &= 0,342422681 \end{aligned}$$

Tabel 4.7 Hasil Perhitungan *idf*

Term	<i>idf</i>
agus	0,740363
ahmad	0,740363
ajar	0,342423
amin	1,041393

asap	0,439333
badan	1,041393
batas	1,041393
bencana	1,041393
..	..
..	..
udara	1,041393
undur	1,041393
wa	1,041393
wibowo	1,041393
widodo	0,740363
zubaida	1,041393
zulinto	0,740363

4.2.4 Perhitungan BM25

Perhitungan BM25 dilakukan sesuai dengan Persamaan 2.2 dengan menentukan nilai parameter k_1 dan b . Nilai k_1 dan b yang digunakan adalah sebesar 1.2 dan b sebesar 0.75 (Manning et al., 2009). Untuk setiap kalimat yang dihitung kepiripannya dengan kalimat lain, maka kalimat tersebut dianggap sebagai *query* yang dihitung nilai kemiripannya dengan kalimat lainnya. Hasil perhitungan BM25 antara kalimat ditunjukkan pada Tabel 4.8.

Contoh perhitungan BM25 antara kalimat 1 dan kalimat 2:

$$RSV_d = \sum_{t \in q} Idf \cdot \frac{(k_1 + 1)tf_{td}}{k_1 \left((1 - b) + b \left(\frac{L_d}{L_{ave}} \right) \right) + tf_{td}}$$

$$RSV_{kalimat1, kalimat2} = \sum_{t \in q} Idf \cdot \frac{(k_1 + 1)tf_{td}}{k_1 \left((1 - b) + b \left(\frac{L_{kalimat2}}{L_{ave}} \right) \right) + tf_{td}}$$

$$score_{sekolah, kalimat2} = 0,56427143 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{sekolah, kalimat2} = 0$$

$$score_{kota, kalimat2} = 0,342422681 \cdot \frac{(1,2 + 1)1}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 1}$$

$$score_{kota, kalimat2} = 0,3564788$$

$$score_{palembang,kalimat2} = 0,263241435 \cdot \frac{(1,2 + 1)2}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 2}$$

$$score_{palembang,kalimat2} = 0,372042461$$

$$score_{sumatera,kalimat2} = 0,740362689 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{sumatera,kalimat2} = 0$$

$$score_{selatan,kalimat2} = 0,740362689 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{selatan,kalimat2} = 0$$

$$score_{putus,kalimat2} = 0,740362689 \cdot \frac{(1,2 + 1)1}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 1}$$

$$score_{putus,kalimat2} = 0,77075386$$

$$score_{pulang,kalimat2} = 0,56427143 \cdot \frac{(1,2 + 1)1}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 1}$$

$$score_{pulang,kalimat2} = 0,587434226$$

$$score_{siswa,kalimat2} = 0,56427143 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{siswa,kalimat2} = 0$$

$$score_{lantar,kalimat2} = 1,041392685 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{lantar,kalimat2} = 0$$

$$score_{kabut,kalimat2} = 0,439332694 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{kabut,kalimat2} = 0$$

$$score_{asap,kalimat2} = 0,439332694 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{asap,kalimat2} = 0$$

$$score_{tebal,kalimat2} = 0,740362689 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{tebal,kalimat2} = 0$$

$$score_{limut,kalimat2} = 1,041392685 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{limut,kalimat2} = 0$$

$$score_{senin,kalimat2} = 1,041392685 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{senin,kalimat2} = 0$$

$$score_{pagi,kalimat2} = 0,740362689 \cdot \frac{(1,2 + 1)0}{1,2 \left((1 - 0,75) + 0,75 \left(\frac{15}{16,6} \right) \right) + 0}$$

$$score_{pagi,kalimat2} = 0$$

$$\begin{aligned} RSV_{kalimat1,kalimat2} &= 0 + 0,3564788 + 0,372042461 + 0 + 0 + 0,77075386 \\ &= +0,587434226 + 0 + 0 + 0 + 0 + 0 + 0 + 0 \\ &= 2,086709347 \end{aligned}$$

Tabel 4.8 Hasil Perhitungan BM25

		Dokumen				
		1	2	3	4	5
Query	1		2,086709	3,935707	1,357244	0
	2	2,020233		2,261879	2,119218	1,131922
	3	4,394107	2,674144		0,702632	0
	4	1,287097	1,999734	0,501789		2,505069
	5	0	1,044801	0	2,576688	
	6	0,260672	0,829409	1,115175	0,305387	1,768339
	7	2,33636	1,271213	1,479159	1,416586	0
	8	2,199502	0	2,279128	0,654612	0
	9	0	0,356479	0	0,397245	0
	10	0,728334	3,631969	1,866365	2,119218	3,239779

Tabel 4.9 Hasil Perhitungan BM25 (Lanjutan)

		Dokumen				
		6	7	8	9	10
Query	1	0,305387	2,456241	1,662959	0	0,523152
	2	0,815057	1,271213	0	0,386204	2,932162

	3	1,684469	1,986118	2,075131	0	2,074821
	4	0,305387	1,271213	0,467495	0,386204	1,577885
	5	1,818895	0	0	0	2,481174
	6		1,071385	0,412171	0	3,133325
	7	1,174798		3,270635	0,386204	1,966903
	8	0,520186	3,527626		0	0,424772
	9	0	0,356479	0		0,935273
	10	4,157976	2,342597	0,412171	1,221227	

4.2.5 Perhitungan *PageRank*

Dalam perhitungan bobot *PageRank*, bobot awal diinisialisasi dengan nilai random antara nol hingga satu, lalu untuk iterasi pertama perhitungan bobot diperbarui menggunakan Persamaan 2.3. Setelah perhitungan bobot *PageRank* untuk semua kalimat dilakukan dalam satu iterasi, bobot *PageRank* diganti dengan nilai baru yang telah dihitung. Contoh inisialisasi bobot awal *PageRank* ditunjukkan pada Tabel 4.10 dan hasil perhitungan bobot *PageRank* untuk dua iterasi ditunjukkan pada Tabel 4.11.

Tabel 4.10 Inisialisasi Awal Bobot *PageRank*

Kalimat	<i>PageRank</i>
1	0,400827866
2	0,863170087
3	0,389187762
4	0,924094751
5	0,157640608
6	0,714980958
7	0,216858534
8	0,237221536
9	0,076112858
10	0,841401681

Contoh perhitungan bobot *PageRank* untuk kalimat 1 iterasi 1:

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

$$WS(V_{kal1}) = (1 - d) + d * \sum_{V_j \in In(V_{kal1})} \frac{w_{ji}}{\sum_{V_k \in Out(V_{kalj})} w_{jk}} WS(V_{kalj})$$

$$score_{kalimat2} = \frac{w_{kal1,kal2}}{\sum_{V_k \in Out(V_{kal2})} w_{kal2,kalk}} WS(V_{kal2})$$

$$score_{kalimat2} = \frac{w_{kal1,kal2}}{(w_{kal2,kal1} + w_{kal2,kal3} + w_{kal2,kal4} + w_{kal2,kal5} + w_{kal2,kal6} + w_{kal2,kal7} + w_{kal2,kal8} + w_{kal2,kal9} + w_{kal2,kal10})} WS(V_{kal2})$$

$$score_{kalimat2} = \frac{2,086709}{(2,020233 + 2,261879 + 2,119218 + 1,131922 + 0,815057 + 1,271213 + 0 + 0,386204 + 2,932162)} 0,863170087$$

$$score_{kalimat2} = 0,101031933$$

$$score_{kalimat3} = \frac{w_{kal1,kal3}}{\sum_{V_k \in Out(V_{kal3})} w_{kal3,kalk}} WS(V_{kal3})$$

$$score_{kalimat3} = \frac{w_{kal1,kal3}}{(w_{kal3,kal1} + w_{kal3,kal2} + w_{kal3,kal4} + w_{kal3,kal5} + w_{kal3,kal6} + w_{kal3,kal7} + w_{kal3,kal8} + w_{kal3,kal9} + w_{kal3,kal10})} WS(V_{kal3})$$

$$score_{kalimat3} = \frac{3,935707}{(4,394107 + 2,674144 + 0,702632 + 0 + 1,684469 + 1,986118 + 2,075131 + 0 + 2,074821)} 0,389187762$$

$$score_{kalimat3} = 0,186310256$$

$$score_{kalimat4} = \frac{w_{kal1,kal4}}{\sum_{V_k \in Out(V_{kal4})} w_{kal4,kalk}} WS(V_{kal4})$$

$$score_{kalimat4} = \frac{w_{kal1,kal4}}{(w_{kal4,kal1} + w_{kal4,kal2} + w_{kal4,kal3} + w_{kal4,kal5} + w_{kal4,kal6} + w_{kal4,kal7} + w_{kal4,kal8} + w_{kal4,kal9} + w_{kal4,kal10})} WS(V_{kal4})$$

$$score_{kalimat4} = \frac{1,357244}{(1,287097 + 1,999734 + 0,501789 + 2,505069 + 0,305387 + 1,271213 + 0,467495 + 0,386204 + 1,577885)} 0,924094751$$

$$score_{kalimat4} = 0,061885647$$

$$score_{kalimat5} = \frac{w_{kal1,kal5}}{\sum_{V_k \in Out(V_{kal5})} w_{kal5,kalk}} WS(V_{kal5})$$

$$score_{kalimat5} = \frac{w_{kal1,kal5}}{(w_{kal5,kal1} + w_{kal5,kal2} + w_{kal5,kal3} + w_{kal5,kal4} + w_{kal5,kal6} + w_{kal5,kal7} + w_{kal5,kal8} + w_{kal5,kal9} + w_{kal5,kal10})} WS(V_{kal5})$$

$$score_{kalimat5} = \frac{0}{(0 + 1,044801 + 0 + 2,576688 + 1,818895 + 0 + 0 + 0 + 2,481174)} 0,157640608$$

$$score_{kalimat5} = 0$$

$$score_{kalimat6} = \frac{w_{kal1,kal6}}{\sum_{V_k \in Out(V_{kal6})} w_{kal6,kalk}} WS(V_{kal6})$$

$$score_{kalimat6} = \frac{w_{kal1,kal6}}{(w_{kal6,kal1} + w_{kal6,kal2} + w_{kal6,kal3} + w_{kal6,kal4} + w_{kal6,kal5} + w_{kal6,kal7} + w_{kal6,kal8} + w_{kal6,kal9} + w_{kal6,kal10})} WS(V_{kal6})$$

$$score_{kalimat6} = \frac{0,305387}{(0,260672 + 0,829409 + 1,115175 + 0,305387 + 1,768339 + 1,071385 + 0,412171 + 0 + 3,133325)} 0,714980958$$

$$score_{kalimat6} = 0,014357623$$

$$score_{kalimat7} = \frac{w_{kal1,kal7}}{\sum_{V_k \in Out(V_{kal7})} w_{kal7,kalk}} WS(V_{kal7})$$

$$score_{kalimat7} = \frac{w_{kal1,kal7}}{(w_{kal7,kal1} + w_{kal7,kal2} + w_{kal7,kal3} + w_{kal7,kal4} + w_{kal7,kal5} + w_{kal7,kal6} + w_{kal7,kal8} + w_{kal7,kal9} + w_{kal7,kal10})} WS(V_{kal7})$$

$$score_{kalimat7} = \frac{2,456241}{(2,33636 + 1,271213 + 1,479159 + 1,416586 + 0 + 1,174798 + 3,270635 + 0,386204 + 1,966903)} 0,216858534$$

$$score_{kalimat7} = 0,121598065$$

$$score_{kalimat8} = \frac{w_{kal1,kal8}}{\sum_{V_k \in Out(V_{kal8})} w_{kal8,kalk}} WS(V_{kal8})$$

$$score_{kalimat8} = \frac{w_{kal1,kal8}}{(w_{kal8,kal1} + w_{kal8,kal2} + w_{kal8,kal3} + w_{kal8,kal4} + w_{kal8,kal5} + w_{kal8,kal6} + w_{kal8,kal7} + w_{kal8,kal9} + w_{kal8,kal10})} WS(V_{kal8})$$

$$score_{kalimat8} = \frac{1,662959}{(2,199502 + 0 + 2,279128 + 0,654612 + 0 + 0,520186 + 3,527626 + 0 + 0,424772)} 0,237221536$$

$$score_{kalimat8} = 0,070786192$$

$$score_{kalimat9} = \frac{w_{kal1,kal9}}{\sum_{V_k \in Out(V_{kal9})} w_{kal9,kalk}} WS(V_{kal9})$$

$$score_{kalimat9} = \frac{w_{kal1,kal9}}{(w_{kal9,kal1} + w_{kal9,kal2} + w_{kal9,kal3} + w_{kal9,kal4} + w_{kal9,kal5} + w_{kal9,kal6} + w_{kal9,kal7} + w_{kal9,kal8} + w_{kal9,kal10})} WS(V_{kal9})$$

$$score_{kalimat9} = \frac{0}{(0 + 0,356479 + 0 + 0,397245 + 0 + 0 + 0,356479 + 0 + 0,935273)} 0,076112858$$

$$score_{kalimat9} = 0$$

$$score_{kalimat10} = \frac{w_{kal1,kal10}}{\sum_{V_k \in Out(V_{kal9})} w_{kal9,kalk}} WS(V_{kal9})$$

$$score_{kalimat10} = \frac{w_{kal1,kal9}}{(w_{kal10,kal1} + w_{kal10,kal2} + w_{kal10,kal3} + w_{kal10,kal4} + w_{kal10,kal5} + w_{kal10,kal6} + w_{kal10,kal7} + w_{kal10,kal8} + w_{kal10,kal9})} WS(V_{kal10})$$

$$score_{kalimat10} = \frac{0,523152}{(0,728334 + 3,631969 + 1,866365 + 2,119218 + 3,239779 + 4,157976 + 2,342597 + 0,412171 + 1,221227)} 0,841401681$$

$$score_{kalimat10} = 0,026420101$$

$$WS(V_{kal1}) = (1 - 0,85) + 0,85 * (0,139217873 + 0,098241767 + 0,121747014 + 0 + 0,024544643 + 0,0400438 + 0,04106776 + 0 + 0,022321948)$$

$$WS(V_{kal1}) = 0,564107084$$

Tabel 4.11 Hasil Perhitungan Bobot *PageRank*

Kalimat	Inisialisasi awal	Iterasi 1	Iterasi 2
1	0,400827866	0,564107084	0,645031343
2	0,863170087	0,626411969	0,727372658
3	0,389187762	0,738073776	0,77323643
4	0,924094751	0,469729805	0,657956046
5	0,157640608	0,619960421	0,472058505
6	0,714980958	0,418234927	0,563115694
7	0,216858534	0,658518382	0,69574379
8	0,237221536	0,408885441	0,540094241
9	0,076112858	0,239364391	0,582118663
10	0,841401681	0,995877088	1,101678699

4.2.6 Mendapatkan Ringkasan

Setelah melakukan pembaruan bobot *PageRank* sebanyak iterasi yang ditentukan, kalimat diurutkan berdasarkan bobot *PageRank*. Jumlah kalimat yang dipilih untuk dijadikan ringkasan ditentukan berdasarkan *compression rate*. Pada Tabel 4.12 ditunjukan jumlah kalimat yang diambil sebanyak 30% dari total panjang dokumen dan dibulatkan keatas. Kalimat yang terpilih kemudian diurutkan kembali berdasarkan urutan aslinya sesuai dengan dokumen awal seperti pada Tabel 4.13.

Tabel 4.12 Hasil Pengurutan Bobot *PageRank*

Kalimat	Iterasi 2
10	1,101679
3	0,773236
2	0,727373

7	0,695744
4	0,657956
1	0,645031
9	0,582119
6	0,563116
8	0,540094
5	0,472059

Tabel 4.13 Hasil Ringkasan Dokumen

Index	Kalimat
2	Kepala SMP Negeri 7 Palembang, Siti Zubaida, mengatakan keputusan pemulangan ditempuh sesuai dengan instruksi Dinas Pendidikan Kota Palembang.
3	"Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan berdampak buruk terhadap siswa, oleh karenanya atas instruksi Kadiknas Kota Palembang melalui pesan WA Grup meminta siswa dipulangkan dan belajar di rumah masing-masing saja," jelas Siti kepada radio Elshinta.
10	"Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud, TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers.

4.3 Perancangan evaluasi

Evaluasi dilakukan dengan membandingkan hasil peringkasan manual yang dilakukan pakar dengan hasil peringkasan yang dihasilkan sistem. Jumlah *compression rate* yang digunakan adalah 10%, 15%, 20%, dan 30%. Rancangan evaluasi dapat dilihat pada Tabel 4.14, Tabel 4.15, Tabel 4.16, Tabel 4.17.

Tabel 4.14 Perancangan Evaluasi Sistem Compression Rate 5%

Dokumen	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Dokumen 1			
..			
..			
Dokumen 10			

Tabel 4.15 Perancangan Evaluasi Sistem Compression Rate 10%

Dokumen	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Dokumen 1			
..			
..			
Dokumen 10			

Tabel 4.16 Perancangan Evaluasi Sistem Compression Rate 20%

Dokumen	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Dokumen 1			
..			
..			
Dokumen 10			

Tabel 4.17 Perancangan Evaluasi Sistem Compression Rate 30%

Dokumen	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Dokumen 1			
..			
..			
Dokumen 10			

BAB 5 IMPLEMENTASI

5.1 Implementasi *Web Scrapping*

Pada pengimplementasian *web scrapping* sistem yang dibuat dibantu dengan menggunakan library BeautifulSoup untuk melakukan *scrapping* pada halaman web yang diinginkan. Selanjutnya data diseleksi untuk mendapatkan bagian judul dan konten berita. Implementasi *web scrapping* ditunjukan pada Kode Sumber 5.1.

Web Scrapping	
1	#bsoup
2	from urllib.request import urlopen
3	from bs4 import BeautifulSoup
4	
5	def get_document(url_bbcnews):
6	
7	try:
8	page_news = urlopen(url_bbcnews)
9	except:
10	return "error", "none"
11	
12	BeautifulSoup = BeautifulSoup(page_news, 'html.parser')
13	title = BeautifulSoup.find('h1', {"class": "story-body__h1"}).text
14	text_news = BeautifulSoup.find('div', {"class": "story-body__inner"})
15	
16	document = ''

Kode Sumber 5.1 Implementasi *Web Scrapping*

Penjelasan dari Kode Sumber 5.1:

Baris 2-3 : Import library BeautifulSoup dan urlopen

Baris 5 : Deklarasi fungsi `get_document` dengan parameter `url_bbcnews`

Baris 7-10 : Melakukan fungsi *try-except* untuk mengakses url, apabila url tidak berhasil diakses maka akan mengembalikan pesan error.

Baris 12 : Mengambil data HTML pada halaman web yang berhasil diakses

Baris 13 : Mengambil bagian judul pada data HTML menggunakan tag `h1`

Baris 14 : Mengambil bagian konten berita pada data HTML menggunakan tag `div`

Baris 16-18 : Bagian konten berita akan diambil data yang berada pada tag `p`, lalu menggabungkan semua konten menjadi satu dokumen

Baris 20 : Mengembalikan konten berita dan judul berita

5.2 Implementasi *Preprocessing*

Pada implementasi *preprocessing* terdapat tahapan segmentasi, *cleaning*, *stemming*, *stopword removal*, dan tokenisasi. Pada tahap ini dibantu dengan *library* NLTK, *pandas*, *Sastrawi*, dan *re*. Tahapan segmentasi menggunakan fungsi `sent_tokenize` dari *library* NLTK. Tahapan *cleaning* dibantu dengan menggunakan *library* *re*. Untuk melakukan penghilangan kata berimbuhan dibantu dengan fungsi `factory.create_stemmer()` dari *library* *Sastrawi*. Sedangkan untuk tahapan tokenisasi menggunakan fungsi `word_tokenize` dari *library* NLTK. Implementasi *preprocessing* ditunjukkan pada Kode Sumber 5.3.

Preprocessing	
1	<code>import pandas as pd</code>
2	<code>import re</code>
3	<code>from nltk.tokenize import sent_tokenize</code>
4	<code>from nltk.tokenize import word_tokenize</code>
5	
6	<code># import StemmerFactory class</code>
7	<code>from Sastrawi.Stemmer.StemmerFactory import StemmerFactory</code>
8	
9	<code># create stemmer</code>
10	<code>factory = StemmerFactory()</code>
11	<code>stemmer = factory.create_stemmer()</code>
12	
13	<code>#this function does cleaning, tokenize, remove stopwords, and</code> <code>stemming</code>
14	<code>def get_clean_corpus(corpus, stopwords):</code>
15	
16	<code> #segmentasi</code>
17	<code> temp = sent_tokenize(corpus)</code>
18	<code> corpus = pd.DataFrame(temp, columns=['dokumen'])</code>
19	
20	<code> clean_corpus = []</code>
21	<code> token = []</code>
22	<code> for index, sentence in enumerate(corpus['dokumen']):</code>
23	<code> term = word_tokenize(corpus['dokumen'][index])</code>
24	
25	<code> #deleting URL</code>
26	<code> deleted_URL = [temp for temp in term if not</code> <code>re.match(r"\w+(?:\.\(\w+)\.\(\w+)\) \w+(?:\.\(\w+)\)", str(temp))]</code>

Kode Sumber 5.2 Implementasi *Preprocessing*

```

27
28         #deleting symbol
29         deleted_symbol = [re.sub(r"[\-
\+=\=:;\\"\\@\[\\\",_!;.':#$$%^&*()<>?/\|}{~:]", " ",str(temp)) for
temp in deleted_URL ]
30
31         #stemming
32         stemmed_sentence = stemmer.stem(" ".join(deleted_symbol))
33
34         tokens = word_tokenize(stemmed_sentence)
35
36         for i in range(len(tokens)):
37             for index, word in enumerate(tokens):
38                 #delete stopwprds
39                 if word in stopwords:
40                     del tokens[index]
41
42                 #delete number
43                 if word.isdigit():
44                     del tokens[index]
45
46         clean_corpus.append(" ".join(tokens))
47         token.append(list(dict.fromkeys(tokens)))
48
49     corpus['clean_corpus'] = clean_corpus
50     corpus['terms'] = token
51
52     return corpus

```

Kode Sumber 5.3 Implementasi *Preprocessing* (Lanjutan)

Penjelasan dari Kode Sumber 5.3:

Baris 1-7 : Melakukan *import library pandas, re, NLTK, dan Sastrawi*

Baris 10-11 : Inisialisasi fungsi *StremmerFactory* dan membuat *object stemmer*

Baris 14 : Deklarasi fungsi *get_clean_corpus* dengan parameter dokumen dan daftar *stopword*

Baris 17 : Segmentasi dokumen

Baris 18 : Inisialisasi *dataframe* corpus dengan argumen hasil segmentasi

Baris 22 : Melakukan perulangan untuk setiap kalimat pada hasil segmentasi

Baris 23 : Melakukan tokenisasi kalimat

Baris 26 : Melakukan penghilangan URL untuk setiap hasil tokenisasi menggunakan *regex* dengan bantuan fungsi *match* dari *library re*

Baris 29 : Melakukan penghilangan simbol pada hasil penghilangan URL

Baris 32 : Melakukan penghilangan imbuhan dengan bantuan object stemmer

Baris 34 : Melakukan tokenisasi kembali dari hasil penghilangan imbuhan

Baris 36-44 : Melakukan perulangan untuk setiap kata dari hasil tokenisasi, apabila kata muncul pada daftar *stopword* maka kata akan dihapus dan apabila kata mengandung angka maka kata juga akan dihapus

Baris 46 : Menyimpan hasil kata yang digabung menjadi dengan pemisahan spasi ke dalam *list* `clean_corpus`

Baris 47 : Menyimpan hasil *list* kata ke dalam *list* token

Baris 49 : Menambahkan kolom `clean_corpus` ke *dataframe* `corpus`

Baris 50 : Menambahkan kolom token ke *dataframe* `corpus`

Baris 52 : Mengembalikan hasil fungsi `get_clean_corpus` dengan *dataframe* `corpus`

5.3 Implementasi *Term Weighting*

Pada implementasi *term weighting* terdapat tahapan pengambilan *term* unik, menghitung *term frequency*, menghitung *document frequency (df)*, dan *inverse document frequency (idf)*. Pada tahap ini dibantu dengan library `pandas` dan `math` dalam proses perhitungan *idf*. Implementasi *term weighting* ditunjukkan pada Kode Sumber 5.5.

Term weighting	
1	<code>import math</code>
2	<code>def get_term_weighting_score(cleaning_result):</code>
3	
4	<code>#getting all the terms</code>
5	<code>terms = []</code>
6	<code>for index, sentence in enumerate(cleaning_result['terms']):</code>
7	<code>terms += [temp for temp in sentence if temp not in terms]</code>
8	<code>terms.sort()</code>
9	
10	<code>#getting frequency for every sentences</code>
11	<code>terms_frequency = pd.DataFrame()</code>
12	<code>for index, term in enumerate(cleaning_result['terms']):</code>
13	<code>frequency_each_sentence = []</code>
14	<code>for i, d in enumerate(terms):</code>
15	<code>temp = term.count(d)</code>

Kode Sumber 5.4 Implementasi *Term Weighting*

```

16         frequency_each_sentence.append(temp)
17         terms_frequency[str(index+1)] = frequency_each_sentence
18
19         terms_frequency['terms'] = terms
20         terms_frequency.set_index('terms', inplace= True)
21
22         #getting df for every terms
23         df_idf = pd.DataFrame(terms_frequency.sum(axis=1),
columns=['df_term'])
24
25         df_idf['terms'] = terms
26         df_idf.set_index('terms', inplace= True)
27
28         #getting idf for every terms
29         N = len(terms_frequency.columns)
30         terms_idf = []
31         for i, d in df_idf.iterrows():
32             idf_score = math.log((N+1)/((df_idf['df_term'][i])), 10)
33             terms_idf.append(idf_score)
34         df_idf['idf_term'] = terms_idf
35
36         return terms_frequency, df_idf

```

Kode Sumber 5.5 Implementasi *Term Weighting* (Lanjutan)

Penjelasan dari Kode Sumber 5.5:

Baris 1 : Melakukan *import library math*

Baris 2 : Deklarasi fungsi *get_term_weighting_score* dengan parameter hasil dari preprocessing

Baris 5-8 : Mendapatkan seluruh *term* unik dari seluruh *list term* pada tiap kalimat dari hasil *preprocessing* lalu mengurutkan term sesuai urutan abjad

Baris 11 : Inisialisasi *dataframe* *terms_frequency*

Baris 12-17 : Untuk setiap *term* dari *term* unik akan dihitung jumlah kemunculan *term* tersebut pada setiap *list term* per kalimat dan menyimpan hasil *tf* pada *dataframe* *terms_frequency*

Baris 19-20 : Menambahkan kolom term unik pada *dataframe* *terms_frequency* dan menjadikan *term* unik sebagai *index dataframe* *terms_frequency*

Baris 23-25 : Menghitung *df* dari *dataframe* *terms_frequency*

Baris 28 : Menghitung jumlah kalimat yang terdapat pada dokumen

Baris 30-33 : Menghitung *idf* dari hasil *df*

Baris 35 : Mengembalikan hasil fungsi `get_clean_corpus` dengan *dataframe* `terms_frequency` dan `df_idf`

5.4 Implementasi Kalimat Sebagai Node

Dalam mengimplementasikan *TextRank* untuk melakukan pruning, dokumen direpresentasikan sebagai graf sedangkan kalimat dalam dokumen direpresentasikan sebagai *node* dan nilai bobot *similarity* BM25 direpresentasikan sebagai *edges* (hubungan antar *node*). Oleh sebab itu diperlukan pembuatan *class* kalimat. Setiap objek kalimat memiliki atribut dan method untuk menghitung bobot *similarity* BM25 dan bobot *PageRank*. Implementasi kalimat sebagai node ditunjukkan pada Kode Sumber 5.7.

Class Sentence	
1	<code>class Sentence:</code>
2	<code> np.random.seed(0)</code>
3	<code> def __init__(self, id, full_sentence, clean_sentence, tokens):</code>
4	<code> self.list_bm25 = {}</code>
5	<code> self.pagerank_score = random.random()</code>
6	<code> self.id = id</code>
7	<code> self.full_sentence = full_sentence</code>
8	<code> self.clean_sentence = clean_sentence</code>
9	<code> self.tokens = tokens</code>
10	<code> self.sentence_len = len(clean_sentence.split())</code>
11	<code> self.pagerank_score_new = 0</code>
12	
13	<code> def calculate_bm25(self, raw_frequency, idf, doc, slen_ave):</code>
14	<code> k1 = 1.2</code>
15	<code> b = 0.75</code>
16	<code> total_bm25 = 0</code>
17	<code> for query in self.tokens:</code>
18	<code> tf = raw_frequency.at[str(query),str(doc.id)]</code>
19	<code> idff = idf.at[str(query),'idf_term']</code>
20	<code> temp = idff * ((k1+1) * tf) / (k1*((1-b) + (b</code>
21	<code> *(doc.sentence_len/slen_ave))) + tf)</code>
22	<code> total_bm25 += temp</code>
23	<code> self.list_bm25[doc.id] = total_bm25</code>
24	<code> def calculate_new_pagerank(self, doc):</code>
25	<code> d=0.85</code>
26	<code> sum_InVi = 0</code>
27	<code> for item in doc:</code>
28	<code> if self.id is not item.id:</code>
	<code> Wji = self.list_bm25[item.id]</code>

Kode Sumber 5.6 Implementasi Kalimat Sebagai Node

29	<code>total_Wjk = sum(item.list_bm25.values())</code>
30	<code>sum_InVi += Wji/total_Wjk*item.pagerank_score</code>
31	<code>self.pagerank_score_new = (1-d)+(d*sum_InVi)</code>

Kode Sumber 5.7 Implementasi Kalimat Sebagai *Node* (Lanjutan)

Penjelasan dari Kode Sumber 5.7:

Baris 1 : Deklarasi *class* Sentence

Baris 2 : Digunakan agar angka random yang pada saat inialisasi bobot awal *PageRank* tidak berubah ketika menjalankan program berulang kali dalam satu waktu

Baris 3 : *Constructor class* Sentence dengan parameter id kalimat, kalimat utuh, kalimat setelah *preprocessing*, dan *term* yang terdapat pada kalimat

Baris 4-11 : Inialisasi nilai atribut ketika melakukan pembuatan objek Sentence

Baris 13 : Deklarasi fungsi *calculate_BM25* dengan parameter *tf*, *idf*, kalimat lain, dan panjang rata-rata dokumen

Baris 14-16 : Inialisasi bobot *k1*, *b*, dan nilai awal *total_BM25*

Baris 17-21 : Menghitung bobot BM25 setiap term dengan kalimat lain lalu menambahkan nilai bobotnya dengan variable *total_BM25*

Baris 22 : Menyimpan nilai kemiripan BM25 kalimat dengan setiap kalimat lainnya

Baris 23 : Deklarasi fungsi *calculate_new_pagerank* dengan parameter seluruh dokumen

Baris 24-27 : Inialisasi bobot *d* dan nilai awal *sum_InVi*

Baris 28-30 : Melakukan perhitungan bobot *sum_InVi* untuk seluruh kalimat pada dokumen selain objek kalimat itu sendiri

Baris 31 : Menyimpan nilai baru *PageRank* dengan nilai $(1-d)+(d*\text{sum_InVi})$

5.5 Implementasi Graf Dokumen

Untuk satu dokumen yang diringkas akan dijadikan objek *graf* dokumen yang memiliki atribut dan beberapa perhitungan untuk mendapatkan ringkasan. Proses ini dimulai dengan pementuan objek Sentence untuk setiap kalimat yang terdapat pada dokumen. Lalu melakukan perhitungan panjang rata-rata kalimat yang digunakan untuk perhitungan BM25. Selanjutnya setiap kalimat dihitung nilai *similarity* BM25 dengan kalimat lainnya. Dalam suatu dokumen tentu saja terjadi kasus yang mana suatu kalimat tidak memiliki nilai *similarity* dengan kalimat manapun, untuk itu dilakukan seleksi kalimat yang disimpan sebagai *outlier*. Kalimat yang sama sekali tidak memiliki bobot *similarity* dengan kalimat lainnya dibuang dari dokumen.

Selanjutnya dilakukan perhitungan bobot *PageRank* sebanyak batas berhenti yang ditentukan. Pada implementasi ini kondisi berhenti dilakukan sebanyak

iterasi. Setelah melakukan update bobot *PageRank*, selanjutnya dilakukan pengambilan ringkasan berdasarkan nilai *PageRank* tertinggi sebanyak *compression rate* yang ditentukan. *Compression rate* digunakan untuk mengambil seberapa banyak total kalimat yang dipilih sebagai ringkasan berdasarkan nilai *PageRank* tertinggi. Selanjutnya ringkasan didapatkan dari kalimat terpilih dan diurutkan kembali sesuai urutannya pada dokumen asli. Implementasi graf dokumen ditunjukkan pada Kode Sumber 5.9.

Class Graph	
1	class Graph:
2	def __init__(self, result_doc, raw_frequency, idf, cr):
3	self.raw_frequency = raw_frequency
4	self.idf = idf.drop(columns=['df_term'])
5	self.result_doc = result_doc
6	self.total_doc = len(result_doc['dokumen'])
7	self.slen_ave = 0
8	self.summarize = []
9	self.doc = []
10	self.outlier = []
11	self.compression_rate = cr
12	
13	for index, item in self.result_doc.iterrows():
14	self.doc.append(Sentence((index+1), item['dokumen'],
15	item['clean_corpus'], item['terms']))
16	
17	#calculate len average
18	temp_len_doc = 0
19	for item in self.doc:
20	temp_len_doc += item.sentence_len
21	self.slen_ave = temp_len_doc/len(self.doc)
22	
23	#calculate bm25 for each object sentence
24	for item in self.doc:
25	for item2 in self.doc:
26	if item.id is not item2.id:
27	item.calculate_bm25(raw_frequency=
28	self.raw_frequency, idf=self.idf, doc=item2,
29	slen_ave=self.slen_ave)
30	
31	##CHECKING IF BM25 SCORE IS 0 (OUTLIER SENTENCE)
32	for index, item in enumerate(self.doc):
33	if sum(item.list_bm25.values()) <= 0:
34	self.outlier.append(self.doc.pop(index))

Kode Sumber 5.8 Implementasi Graf Dokumen

```

32
33         #calculate pagerank
34         for i in range(100):
35             for item in self.doc:
36                 item.calculate_new_pagerank(self.doc)
37
38             #update pagerank score
39             for item in self.doc:
40                 item.pagerank_score = item.pagerank_score_new
41
42             #getting the summarize
43             sorted_doc = sorted(self.doc, key=lambda x:
44 x.pagerank_score, reverse=True)
45             top_pagerank = []
46             for item in
47 range(math.ceil(self.total_doc*self.compression_rate)):
48                 top_pagerank.append(sorted_doc[item])
49
50             sorted_sum = sorted(top_pagerank, key=lambda x: x.id)
51
52             temp_summarize = [item.full_sentence for item in
53 sorted_sum]
54
55             self.summarize = temp_summarize

```

Kode Sumber 5.9 Implementasi Graf Dokumen (Lanjutan)

Penjelasan dari Kode Sumber 5.9:

Baris 1 : Deklarasi *class* Graph

Baris 2 : *Constructor class* Graph dengan parameter dokumen, *terms frequency*, *df&idf*, dan bobot *compression rate*

Baris 3-11 : Inisialisasi nilai atribut ketika melakukan pembuatan objek Graph

Baris 13-14 : Untuk setiap kalimat yang terdapat dalam dokumen dilakukan pembuatan objek Sentence dan disimpan sebagai atribut doc (*list of object*)

Baris 17-20 : Menghitung rata-rata panjang kalimat dan menyimpan nilai tersebut sebagai atribut *slen_ave*

Baris 23-26 : Menghitung BM25 setiap objek Sentence dengan memanggil fungsi *calculate_bm25*

Baris 29-31 : Melakukan seleksi kalimat ketika tidak memiliki nilai *similarity* sama sekali dengan kalimat lain. Objek Sentence tersebut akan disimpan pada atribut *outlier*

Baris 34-40 : Melakukan perhitungan bobot *PageRank* untuk sisa kalimat yang telah diseleksi dengan memanggil fungsi *calculate_new_pagerank* dan melakukan update bobot *PageRank* dengan bobot baru sebanyak iterasi yang ditentukan

Baris 43 : Mengurutkan kalimat berdasarkan nilai bobot *PageRank* terbesar

Baris 44-46 : Kalimat dengan bobot *PageRank* terbesar akan dipilih sebanyak *compression rate* dan disimpan sebagai *top_pagerank*

Baris 48 : Mengurutkan kalimat terpilih berdasarkan urutan idnya (sesuai urutannya pada dokumen asli) dan disimpan sebagai *sorted_sum*

Baris 50-52 : Mengambil kalimat utuh dari hasil *sorted_sum* dan menyimpannya sebagai atribut *summarize*

5.6 Implementasi *Main* Sistem Peringkasan

Sistem peringkasan dimulai dengan melakukan *import* URL berita dan nilai *compression rate*. Selanjutnya dilakukan *import* dokumen *stopword* tala yang digunakan pada *preprocessing*. Proses dilakukan dengan *preprocessing* dan perhitungan *term weighting*. Percobaan pada sistem yang dibuat bergantung pada bobot *compression rate* yang diujikan. Percobaan pada penelitian ini menggunakan bobot *compression rate* sebanyak empat jenis, yaitu 5%, 10%, 20%, dan 30%. Main sistem yang dibuat pada penelitian ini ditunjukkan pada Kode Sumber 5.11.

Main system peringkasan	
1	from IPython.display import Markdown, display
2	document2 = input("URL berita BBC Indonesia: ")
3	
4	print("Jenis compression rate:\n 1. 5% \n 2. 10% \n 3. 20% \n 4. 30%")
5	cr2 = input("Pilih jenis compression rate: ")
6	
7	
8	if cr2 in ["1","2","3","4"] :
9	stopword = open("../stopword_list_tala.txt", "r")
10	stopwords = stopword.read().split("\n")
11	document3, title = get_document(document2)
12	
13	if document3 == "error" or title == "none":
14	display(Markdown('***sorry, i cant access the url**'))
15	else:
16	cleaning_result2 = get_clean_corpus(corpus=document3,
17	stopwords=stopwords)
18	terms_frequency2, df_idf2 =
19	get_term_weighting_score(cleaning_result=cleaning_result2)
20	if cr2 == "1":
	crate = 0.05

Kode Sumber 5.10 Main Sistem

```

21         elif cr2 == "2":
22             crate = 0.1
23         elif cr2 == "3":
24             crate = 0.2
25         elif cr2 == "4":
26             crate = 0.3
27
28         percobaan2 = Graph(result_doc=cleaning_result2,
29 raw_frequency=terms_frequency2,
30                             idf=df_idf2, cr=float(crate))
31         display(Markdown('**Judul berita:**'))
32         print(title)
33
34         summarize = ""
35         for ringkasan in percobaan2.summarize:
36             summarize += ringkasan + " "
37
38         display(Markdown('**Hasil ringkasan:**'))
39         print(summarize)
40     else:
41         display(Markdown('**sorry, wrong input**'))

```

Kode Sumber 5.11 Main Sistem (Lanjutan)

Penjelasan dari Kode Sumber 5.11:

Baris 1 : Melakukan *import library* IPython.display

Baris 2 : Melakukan input URL dan *compression rate* oleh user

Baris 4-5 : Menampilkan jenis *compression rate* yang digunakan pada sistem dan meminta *user* untuk memilih *compression rate* yang diinginkan

Baris 8 : Melakukan pengecekan *compression rate* pilihan yang dimasukan *user*

Baris 9-10 : Melakukan import *stopword* tala dalam *file txt* dan menyimpannya dengan nama *variable* stopwords

Baris 11 : Memanggil fungsi *get_documen* dengan parameter url yang telah diberikan dan menyimpan hasil kembaliannya sebagai *variable* document3 dan title

Baris 13 : Melakukan pengecekan nilai *variable* document3 dan title

Baris 14 : Apabila pengecekan bernilai benar maka akan mengembalikan pesan error

Baris 15 : Apabila pengecekan bernilai salah maka melanjutkan ke baris berikutnya

Baris 16 : Memanggil fungsi *get_clean_corpus* dengan parameter stopwords dan document untuk *preprocessing* dan menyimpan hasil kembaliannya sebagai *variable* cleaning_result2

Baris 17 : Memanggil fungsi `get_term_weighting_score` dengan parameter `cleaning_result` untuk *term weighting* dan menyimpan hasil kembaliannya sebagai *variable* `terms_frequency2` dan `df_idf2`

Baris 19-26 : Menentukan nilai *variable* `crate` sesuai dengan jenis *compression rate* yang dimasukkan *user*

Baris 28 : Membuat objek graf dengan parameter dokumen, *terms frequency*, *df&idf*, dan bobot *compression rate* dan menyimpan objek tersebut dengan nama `percobaan2`

Baris 30 : Mencetak judul berita

Baris 33-35 : Menggabungkan hasil ringkasan untuk setiap kalimat yang terdapat pada atribut `summarize` dari objek `percobaan2`

Baris 38 : Mencetak ringkasan berita

Baris 39-40 : Apabila jenis *compression rate* yang dimasukkan *user* tidak sesuai dengan jenis yang disediakan akan mencetak pesan *error*

5.7 Hasil Implementasi

Sesuai dengan implementasi dengan menggunakan bahasa pemrograman python, hasil ringkasan otomatis yang didapatkan sistem ditunjukkan pada Gambar 5.1. Hasil peringkasan tersebut menampilkan ringkasan sesuai urutan dokumen asli berdasarkan URL berita dan *compression rate* yang dimasukkan *user*.

```
URL berita BBC Indonesia: https://www.bbc.com/indonesia/indonesia-48102658
Jenis compression rate:
1. 5%
2. 10%
3. 20%
4. 30%
Pilih jenis compression rate: 4

Judul berita:

Banjir Bengkulu: Dianggap terparah hingga tudingan 'akibat tambang dan sawit'

Hasil ringkasan:

Kerusakan lingkungan disebut Badan Nasional Penanggulangan Bencana (BNPB) sebagai salah satu pemicu banjir dan longsor di Bengkulu. Walau terdapat fakta pembukaan pertambangan batu bara dan perkebunan kelapa sawit di beberapa titik di provinsi itu, sejumlah kalangan menyangkal aktivitas tersebut memicu banjir. Sirajon, warga Desa Arga Mulya, Bengkulu Selatan, mengambil langkah seribu saat arus deras air masuk ke rumahnya yang berjarak 50 meter dari bibir sungai, akhir pekan lalu. "Dampak hujan mungkin karena intensitas hujan tinggi, tapi di sini juga ada perkebunan sawit sehingga hutan untuk menghambat air sudah tidak seperti dulu lagi," tuturnya. Banjir ini juga menimpa Yessi Ameilia Safitri, warga Kelurahan Tanjung Agung, yang berjarak sekitar 230 kilometer dari Arga Mulya. Namun ia menggarisbawahi pula kerusakan lingkungan di sekitar sungai yang memperburuk dampak hujan besar. Kantor berita Antara menyebut delapan pemegang izin batu bara di Bengkulu itu adalah PT Bengkulu Bio Energi, PT Kusuma Ray a Utama, PT Bara Mega Quantum, PT Inti Bara Perdana. Ada pula satu izin perkebunan kelapa sawit yang dikeluarkan untuk PT Agria ndalas yang berada di daerah tangkapan air Sungai Bengkulu. Di sisi lain, pejabat Kementerian Lingkungan Hidup dan Kehutanan (KLHK) menyangkal aktivitas beberapa perusahaan ini menurunkan kemampuan air tanah menyerap air hujan. Jadi banjir ini adalah limpasan, curah hujan melebihi kapasitas sungai," kata Kasubdit Monitoring dan Evaluasi DAS KLHK, Ernawati. Bagaimanapun, dampak banjir dan longsor di Bengkulu masih terus dirasakan warga setempat, meski air berangsur surut. Badan Meteorologi, Klimatologi, dan Geofisika memperkirakan curah hujan tinggi masih akan melanda Bengkulu dan beberapa daerah lain seperti Aceh, Sumatera Selatan, Sulawesi Utara, Papua, dan Papua Barat.
```

Gambar 5.1 Contoh Hasil Peringkasan Sistem

BAB 6 PENGUJIAN DAN ANALISIS

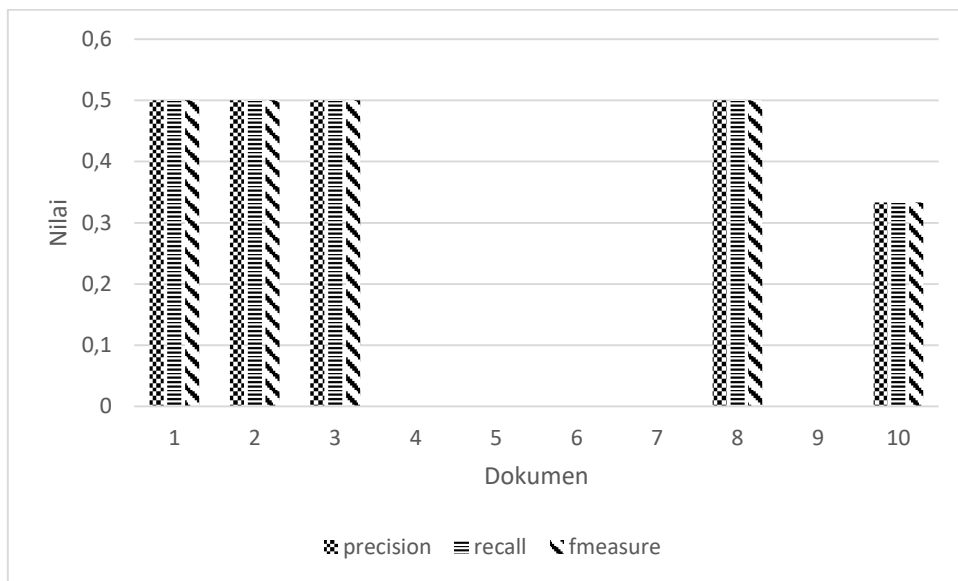
Pengujian dilakukan dengan membandingkan hasil ringkasan 10 artikel berita yang didapatkan dari pakar dan hasil ringkasan yang didapatkan dari sistem. Ringkasan dihasilkan berdasarkan bobot *compression rate* yang digunakan. Pada evaluasi ini bobot yang digunakan adalah 5%, 10%, 20%, dan 30%. Tingkat kualitas yang didapatkan dihitung sesuai perhitungan *precision*, *recall*, dan *f-measure*.

6.1 Hasil Pengujian dan Analisis *Compression Rate* 5%

Tabel 6.1 menunjukkan nilai dari *precision*, *recall*, dan *f-measure* dari sepuluh dokumen yang diuji menggunakan nilai *compression rate* sebesar 5%. Beberapa dokumen memiliki nilai *precision*, *recall*, dan *f-measure* sebesar nol, yang mana menunjukkan bahwa tidak ada kalimat ringkasan yang sama antara hasil ringkasan sistem otomatis dengan hasil ringkasan yang didapatkan oleh pakar. Untuk 10 dokumen yang diuji nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,233; 0,233; dan 0,233. Perbandingan nilai *precision*, *recall*, dan *f-measure* untuk setiap dokumen dapat dilihat pada Gambar 6.1.

Tabel 6.1 Hasil Pengujian *Compression Rate* 5%

Dokumen	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Dokumen 1	0,5	0,5	0,5
Dokumen 2	0,5	0,5	0,5
Dokumen 3	0,5	0,5	0,5
Dokumen 4	0	0	0
Dokumen 5	0	0	0
Dokumen 6	0	0	0
Dokumen 7	0	0	0
Dokumen 8	0,5	0,5	0,5
Dokumen 9	0	0	0
Dokumen 10	0,33	0,33	0,33
Rata - rata	0,233	0,233	0,233



Gambar 6.1 Evaluasi Pengujian *Compression Rate* 5%

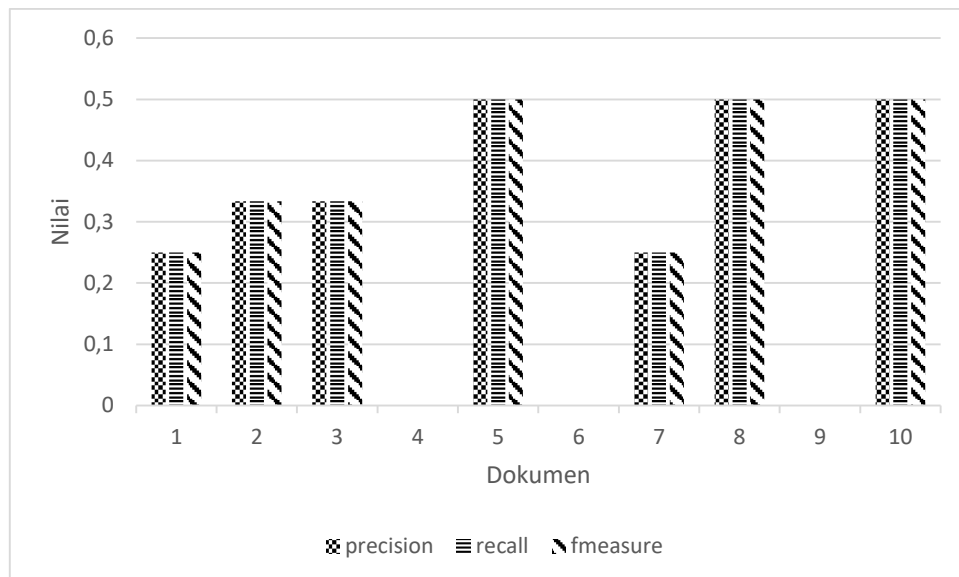
6.2 Hasil Pengujian dan Analisis *Compression Rate* 10%

Tabel 6.2 menunjukkan nilai dari *precision*, *recall*, dan *f-measure* dari sepuluh dokumen yang diuji menggunakan nilai *compression rate* sebesar 10%. Berbeda dengan Tabel 6.1, pengujian dengan menggunakan *compression rate* 10% memiliki lebih sedikit nilai *precision*, *recall*, dan *f-measure* sebesar nol. Hal ini menunjukkan bahwa hasil ringkasan yang didapatkan sistem otomatis memiliki lebih banyak kesamaan dengan hasil ringkasan yang didapatkan pakar dibandingkan dengan pengujian dengan menggunakan *compression rate* 5%. Untuk 10 dokumen yang diuji nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,267; 0,267; dan 0,267. Perbandingan nilai *precision*, *recall*, dan *f-measure* untuk setiap dokumen dapat dilihat pada Gambar 6.2.

Tabel 6.2 Hasil Pengujian *Compression Rate* 10%

Dokumen	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Dokumen 1	0,25	0,25	0,25
Dokumen 2	0,33	0,33	0,33
Dokumen 3	0,33	0,33	0,33
Dokumen 4	0	0	0
Dokumen 5	0,5	0,5	0,5
Dokumen 6	0	0	0
Dokumen 7	0,25	0,25	0,25
Dokumen 8	0,5	0,5	0,5
Dokumen 9	0	0	0

Dokumen 10	0,5	0,5	0,5
Rata-rata	0,267	0,267	0,267



Gambar 6.2 Evaluasi Pengujian *Compression Rate* 10%

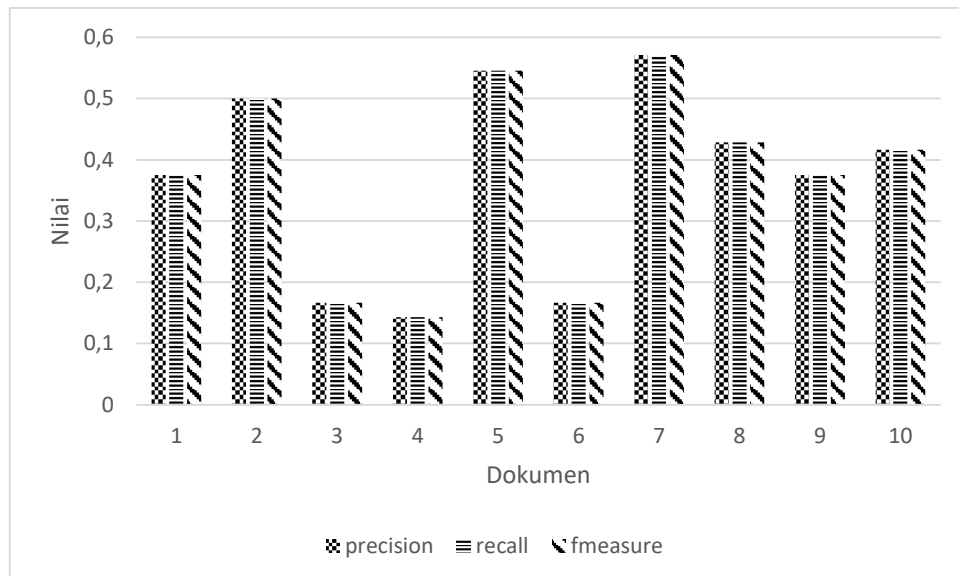
6.3 Hasil Pengujian dan Analisis *Compression Rate* 20%

Tabel 6.3 menunjukkan nilai dari *precision*, *recall*, dan *f-measure* dari sepuluh dokumen yang diuji menggunakan nilai *compression rate* sebesar 20%. Untuk pengujian ini tidak lagi mendapatkan hasil nilai *precision*, *recall*, dan *f-measure* sebesar nol. Hasil *precision*, *recall*, dan *f-measure* terendah didapatkan pada dokumen 3, dokumen 4, dan dokumen 6. Sedangkan untuk hasil *precision*, *recall*, dan *f-measure* tertinggi didapatkan pada dokumen 7 yaitu sebesar 0,571429; 0,571429; dan 0,571429. Untuk 10 dokumen yang diuji nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,368831; 0,368831; dan 0,368831. Perbandingan nilai *precision*, *recall*, dan *f-measure* untuk setiap dokumen dapat dilihat pada Gambar 6.3.

Tabel 6.3 Hasil Pengujian *Compression Rate* 20%

Dokumen	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Dokumen 1	0,375	0,375	0,375
Dokumen 2	0,5	0,5	0,5
Dokumen 3	0,166667	0,166667	0,166667
Dokumen 4	0,142857	0,142857	0,142857
Dokumen 5	0,545455	0,545455	0,545455
Dokumen 6	0,166667	0,166667	0,166667

Dokumen 7	0,571429	0,571429	0,571429
Dokumen 8	0,428571	0,428571	0,428571
Dokumen 9	0,375	0,375	0,375
Dokumen 10	0,416667	0,416667	0,416667
Rata-rata	0,368831	0,368831	0,368831



Gambar 6.3 Evaluasi Pengujian *Compression Rate* 20%

6.4 Hasil Pengujian dan Analisis *Compression Rate* 30%

Tabel 6.4 menunjukkan nilai dari *precision*, *recall*, dan *f-measure* dari sepuluh dokumen yang diuji menggunakan nilai *compression rate* sebesar 30%. Hasil *precision*, *recall*, dan *f-measure* terendah didapatkan pada dokumen 1, yaitu sebesar 0,416667; 0,416667; dan 0,416667. Sedangkan untuk hasil *precision*, *recall*, dan *f-measure* tertinggi didapatkan pada dokumen 7 yaitu sebesar 0,7; 0,7; dan 0,7. Untuk 10 dokumen yang diuji nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,551692; 0,551692; dan 0,551692. Perbandingan nilai *precision*, *recall*, dan *f-measure* untuk setiap dokumen dapat dilihat pada Gambar 6.5.

Contoh perbedaan hasil ringkasan dokumen 7 yang didapatkan pakar dan sistem dapat dilihat pada Gambar 6.4. Kalimat terakhir pada gambar tersebut dipilih pakar sebagai ringkasan karna merupakan kalimat penjelas untuk kalimat sebelumnya, namun kalimat tersebut tidak dipilih oleh sistem. Hal ini terjadi karna jumlah kata yang sedikit pada sebuah kalimat akan mempengaruhi besar nilai *similarity* antara kalimat tersebut dengan kalimat lain. Hal tersebut juga dipengaruhi oleh proses penghilangan kata *stopword* dan proses *stemming*.

Kalimat yang memiliki sedikit nilai *similarity* dengan kalimat lain akan memiliki nilai *PageRank* yang rendah, sehingga tidak akan terpilih untuk dijadikan ringkasan.

Menurutnya, kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan, kegiatan belajar mengajar di tingkat TK sampai SMP akan tetap diliburkan.
Akan tetapi, sebagaimana dipaparkan Kepala Dinas Pendidikan Sumatera Selatan, Widodo, kegiatan belajar mengajar di daerah yang tidak terdampak kabut asap tetap berlangsung.
"Daerah yang tidak terdampak kabut asap tetap normal tetap belajar, untuk daerah yang terkategori sedang tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker, bagi daerah terkategori parah maka siswa diberikan tugas dengan memaksimalkan kelas daring," kata Widodo kepada Antara.
Hal itu belakangan dibenarkan Agus Wibowo, selaku Kepala Pusat Data, Informasi dan Humas Badan Nasional Penanggulangan Bencana (BNPB).
"Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud, TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers.
Sejumlah warga Palembang, Sumatera Selatan, mengeluhkan kabut asap pada Senin (14/10) tergolong parah.
Bahkan jarak pandang hanya 10 meter.

Gambar 6.4 Contoh Hasil Kalimat Yang Terpilih Sebagai Ringkasan

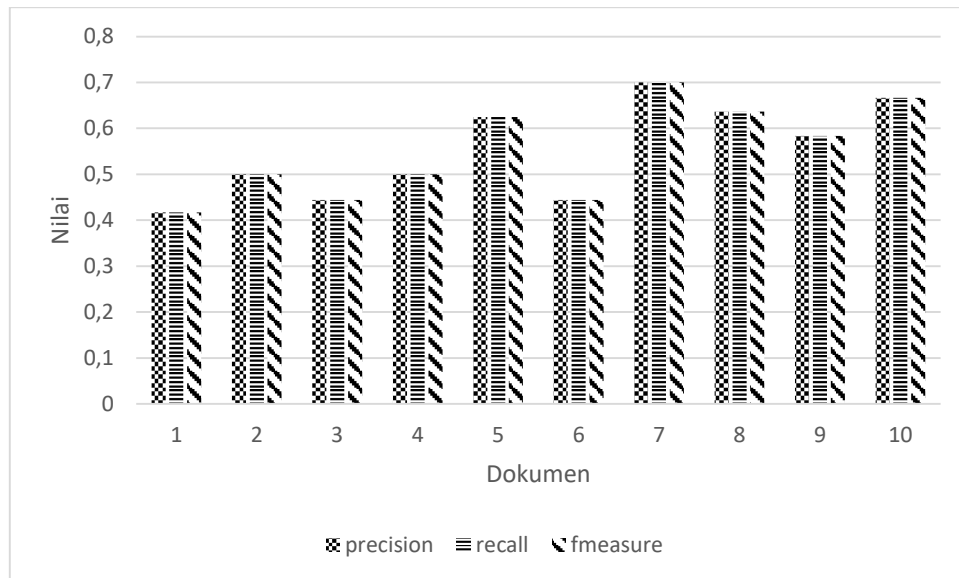
Keterangan :

- : Kalimat terpilih sebagai ringkasan oleh pakar
- : Kalimat terpilih sebagai ringkasan oleh sistem
- : Kalimat terpilih sebagai ringkasan oleh sistem dan pakar

Tabel 6.4 Hasil Pengujian *Compression Rate* 30%

Dokumen	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Dokumen 1	0,416667	0,416667	0,416667
Dokumen 2	0,5	0,5	0,5
Dokumen 3	0,44	0,44	0,44
Dokumen 4	0,5	0,5	0,5
Dokumen 5	0,625	0,625	0,625
Dokumen 6	0,44	0,44	0,44
Dokumen 7	0,7	0,7	0,7
Dokumen 8	0,636364	0,636364	0,636364

Dokumen 9	0,583333	0,583333	0,583333
Dokumen 10	0,667	0,667	0,667
Rata-rata	0,551692	0,551692	0,551692



Gambar 6.5 Evaluasi Pengujian *Compression Rate* 30%

BAB 7 PENUTUP

7.1 Kesimpulan

Berdasarkan implementasi dan evaluasi sistem peringkasan otomatis yang telah berhasil dilakukan dapat disimpulkan beberapa hal sebagai berikut:

1. Sistem peringkasan artikel berita berbahasa Indonesia menggunakan TextRank dengan pembobotan BM25 dapat menghasilkan ringkasan secara ekstraktif dengan cara memeringkatkan setiap kalimat pada satu artikel berita. Jumlah kalimat yang dijadikan ringkasan dipengaruhi oleh besar *compression rate* yang digunakan.
2. Setelah membandingkan hasil ringkasan yang didapatkan sistem peringkasan otomatis dengan hasil ringkasan yang didapatkan dari *expert* (pakar) sebanyak 10 dokumen, kualitas ringkasan terbaik didapatkan pada saat penggunaan *compression rate* sebesar 30% dengan nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,551692; 0,551692; dan 0,551692.

7.2 Saran

Beberapa hal yang dapat diperbaiki dan menjadi masukan untuk penelitian selanjutnya ialah agar dapat memperhatikan posisi kalimat pada saat melakukan peringkasan ekstraktif, memperhatikan kutipan kalimat yang didapatkan dari narasumber pada artikel berita, dan memperhatikan kalimat yang dicetak miring atau dicetak tebal.

DAFTAR RUJUKAN

- Abbasi-ghalehtaki, R., Khotanlou, H. and Esmailpour, M., 2016. Fuzzy evolutionary cellular learning automata model for text summarization. *Swarm and Evolutionary Computation*, [online] 30, pp.11–26. Available at: <<http://dx.doi.org/10.1016/j.swevo.2016.03.004>>.
- Alguliev, R. and Aliguliyev, R., 2009. Evolutionary Algorithm for Extractive Text Summarization. 2009(November), pp.128–138.
- Anon 2020. *Demo Page for TextRank Algorithm*. [online] Available at: <https://publicb2.ceshine.net/file/ceshine-public/misc/textrank_demo.html?source=post_page-----b8efc7e70b4-----> [Accessed 28 Feb. 2020].
- Barrios, F., López, F., Argerich, L. and Wachenchauser, R., 2016. Variations of the Similarity Function of TextRank for Automated Summarization. [online] Available at: <<http://arxiv.org/abs/1602.03606>>.
- Boeing, G. and Waddell, P., 2017. New Insights into Rental Housing Markets across the United States: Web Scraping and Analyzing Craigslist Rental Listings. *Journal of Planning Education and Research*, 37(4), pp.457–476.
- Brin, S. and Page, L., 1998. The anatomy of a large-scale hypertextual Web search engine BT - Computer Networks and ISDN Systems. *Computer Networks and ISDN Systems*, [online] 30(1–7), pp.107–117. Available at: <[http://dx.doi.org/10.1016/S0169-7552\(98\)00110-X%5Cnhttp://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=6&SID=X1pOWPMuSmOv1SlwJ6f&page=1&doc=2](http://dx.doi.org/10.1016/S0169-7552(98)00110-X%5Cnhttp://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=6&SID=X1pOWPMuSmOv1SlwJ6f&page=1&doc=2)>.
- Fang, C., Mu, D., Deng, Z. and Wu, Z., 2017. Word-sentence co-ranking for automatic extractive text summarization. *Expert Systems with Applications*, [online] 72, pp.189–195. Available at: <<http://dx.doi.org/10.1016/j.eswa.2016.12.021>>.
- Fhadli, M., Fauzi, M.A. and Afirianto, T., 2017. Peringkasan Literatur Ilmu Komputer Bahasa Indonesia Berbasis Fitur Statistik dan Linguistik menggunakan Metode Gaussian Naïve Bayes. 1(4), pp.307–319.
- Hahn, U. and Mani, I., 2000. of Automatic Researchers are investigating summarization tools and methods that. *Computer* 33.11, (November), pp.29–36.
- Lv, Y. and Zhai, C., 2011. Lower-bounding term frequency normalization. *International Conference on Information and Knowledge Management, Proceedings*, pp.7–16.
- Manning, C.C.; Ragghavan, P.; Schütze, H., 2009. *An Introduction to Information Retrieval*. Cambridge, England: Cambridge University Press.

- Munot, N. and S. Govilkar, S., 2013. Comparative Study of Text Summarization in Indian Languages. *International Journal of Computer Applications*, 75(6), pp.17–21.
- Mussina, A., Aubakirov, S. and Trigo, P., 2018. Automatic Document Summarization based on Statistical Information. (Data), pp.71–76.
- Niu, J., Zhao, Q., Wang, L., Chen, H., Atiquzzaman, M. and Peng, F., 2016. OnSeS: A novel online short text summarization based on BM25 and neural network. *2016 IEEE Global Communications Conference, GLOBECOM 2016 - Proceedings*, pp.1–6.
- Pinandhita, R.R., 2013. Peringkat Dokumen Berbahasa Indonesia Berbasis Kata Benda Dengan BM25.
- Radev, D. R., Hovy, E., & McKeown, K., 2002. Introduction to the special issue on summarization. *Computational Linguistics*, 28(4), pp.399–408.
- Sankarasubramaniam, Y., Ramanathan, K. and Ghosh, S., 2014. Text summarization using Wikipedia. *Information Processing and Management*, [online] 50(3), pp.443–461. Available at: <<http://dx.doi.org/10.1016/j.ipm.2014.02.001>>.
- Steinberger, J. and Ježek, K., 2009. Evaluation measures for text summarization. *Computing and Informatics*, 28(2), pp.251–275.
- Tarau, R.M. and P., 1973. TextRank: Bringing Order into Texts. *Comparative Biochemistry and Physiology -- Part B: Biochemistry and*, [online] 45(4). Available at: <<http://www.aclweb.org/anthology/W04-3252>>.
- Yeh, J.Y., Ke, H.R., Yang, W.P. and Meng, I.H., 2005. Text summarization using a trainable summarizer and latent semantic analysis. *Information Processing and Management*, 41(1), pp.75–95.

LAMPIRAN A SURAT PERNYATAAN PARTISIPASI PAKAR

	KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI	
	UNIVERSITAS BRAWIJAYA	
	FAKULTAS ILMU KOMPUTER	
	Jl. Veteran No.8, Malang, 65145, Indonesia Telp. : +62-341-577911; Fax : +62-341-577911 http://filkom.ub.ac.id E-mail : filkom@ub.ac.id	

Nomor	: 2515 /UN10.F15.05/PP/2019	15 NOV 2019
Perihal	: <i>Pemohonan menjadi pakar untuk kebutuhan skripsi</i>	

Yth. Prima Zulvarina, S.S., M.Pd.
Jl. Veteran No. 8
Malang

Untuk mendukung penyelesaian skripsi mahasiswa berikut :

Nama	: Yurdha Fadhila Hernawan
NIM	: 165150200111094
Judul Skripsi	: Peringkasan Artikel Berbahasa Indonesia Menggunakan TextRank Dengan Pembobotan BM25
Dosen Pembimbing	: 1. Putra Pandu Adikara, S.Kom, M.Kom 2. Randy Cahya Wihandika, S.St, M.Kom.
Prodi	: Teknik Informatika

Guna mengajukan kesediaan anda selaku Pakar dalam pengerjaan skripsi mahasiswa tersebut.

Atas perhatian dan kerjasamanya kami ucapkan terima kasih.

Plt. Kepala Tata Usaha,



Heru Dintanufi, SE., M.Si.
NP168012262005021001

Tembusan Kepada Yth:

1. Ketua Program Studi Teknik Informatika
2. Mahasiswa yang bersangkutan

LAMPIRAN B HASIL KOMPARASI RINGKASAN SISTEM DAN RINGKASAN PAKAR

B.1 Hasil ringkasan artikel berita untuk evaluasi terburuk dengan *compression rate* 5%

Artikel asli berita 4:

Sejumlah kabupaten dan kota di Sumatera Barat dilanda banjir seiring hujan lebat yang terjadi sejak Kamis (11/10).
Bencana alam ini dilaporkan menyebabkan enam orang tewas dan ratusan rumah terendam air dengan kedalaman hingga satu meter.
Merujuk data Pusat Pengendalian Operasi Penanggulangan Bencana (Pusdalops) BPBD Sumbar, tiga korban tewas berada di Padang Pariaman, sementara tiga lainnya di Tanah Datar.
Adapun, daerah terdampak banjir adalah Tanah Datar, Limapuluh Kota, Pesisir Selatan, Pasaman, Pasaman Barat, Padang Pariaman, Agam, Mentawai Sijunjung, Solok dan Sawahlunto.
Kepala BPBD Sumbar, Erman Rahman, menyebut hujan lebat tak hanya memicu banjir, tapi juga longsor di sejumlah titik, satu di antaranya di Pessel, Bukit Buah Palo.
"Hingga saat ini kami masih mendata korban banjir, tanah longsor dan jembatan putus," ujar Rahman, Jumat (12/10), kepada wartawan di Padang, Agusmanto, yang melapor untuk BBC News Indonesia.
Banjir yang terjadi membuat sebuah jembatan sepanjang delapan meter menuju Benteng Gunung Tak Jadi amblas.
Jembatan itu berada di Kecamatan Bonjol, Kabupaten Pasaman.
Camat Bonjol, Yohanis, menyebut jembatan tersebut tidak bisa dipakai lagi.
Ia berkata, banjir juga membuat jalan di sekitar jembatan rusak.
"Jembatan ini salah satu akses bagi pengendara sepeda motor dari arah pasar Bonjol menuju benteng dan perkapungan warga," kata Yohanis.
Yohanis menyebut jembatan itu berpotensi hanyut seiring banjir yang tak kunjung surut.
"Kami mengimbau warga Bonjol dan sekitarnya tidak melewati jembatan tersebut," ucapnya.
Jalan lintas Sumatera di Kelurahan Berangin, Bonjol, juga dilaporkan sempat terputus akibat longsor dan pohon kayu yang bertumbangan ke jalan.
Polisi kini menerapkan sistem buka-tutup di jalan itu.
Kepala BPBD Pasaman, Masfet Kenedi menuturkan, material longsor yang menimbun jalan di Berangin mencapai lebih dari setengah meter.

Pohon kayu sepanjang 18 meter juga tumbang ke jalan.

Sebuah mobil ambulans dilaporkan terjebak material longsor di Berangin, di jalur yang menghubungkan Bukittinggi dengan Sumatera Utara.

Mobil itu dalam posisi menanjak, lantas terjebak material longsor dari arah kiri.

Kepala Dinas Kesehatan Pasaman, Amdarisman, mengatakan, semua penumpang di mobil tersebut selamat, termasuk ibu hamil yang ada di dalam ambulans tersebut.

"Tidak ada korban jiwa.

Pasien (ibu hamil) sudah dirujuk ke RSUD Pasaman," ujarnya.

Longsor yang terjadi Ranah Batu, Jorong Piubuah, Nagari Tanjung Bonai, Kecamatan Lintau Buo Utara, Tanah Datar terjadi Kamis kemarin, sekitar pukul 18.30 WIB.

Dua orang dilaporkan tewas tertimbun material longsor, sementara dua orang lainnya lainnya hingga kini masih dalam pencarian.

Wali Nagari Jorong Piubuah, Haris, menyebut longsor menghancurkan setengah kampungnya.

Harie berkata, longsor terjadi secara cepat sehingga para sejumlah warga tak sempat menyelamatkan diri.

Longsor berasal dari Gunung Sago yang kemudian dengan cepat menyapu perkampungan di lembahnya.

"Ada dua korban, salah satunya ditemukan 2 kilometer dari rumahnya.


Korban ini sedang hamil," kata Haris.


Kerusakan di lembah Gunung Sago itu mencapai panjang dua kilometer dan sementara kedalaman lumpur berkisar 1,5 meter.

"Total korban terdampak sebanyak 26 jiwa, rata-rata terluka dan sudah dievakuasi ke rumah sakit," ujar Haris.

Kalimat yang dipilih sebagai ringkasan pada artikel berita 4:

Keterangan :

 : Kalimat terpilih sebagai ringkasan oleh pakar

 : Kalimat terpilih sebagai ringkasan oleh sistem

Sejumlah kabupaten dan kota di Sumatera Barat dilanda banjir seiring hujan lebat yang terjadi sejak Kamis (11/10).

Bencana alam ini dilaporkan menyebabkan enam orang tewas dan ratusan rumah terendam air dengan kedalaman hingga satu meter.

Merujuk data Pusat Pengendalian Operasi Penanggulangan Bencana (Pusdalops) BPBD Sumbar, tiga korban tewas berada di Padang Pariaman, sementara tiga lainnya di Tanah Datar.

Adapun, daerah terdampak banjir adalah Tanah Datar, Limapuluh Kota, Pesisir Selatan, Pasaman, Pasaman Barat, Padang Pariaman, Agam, Mentawai Sijunjung, Solok dan Sawahlunto.

Kepala BPBD Sumbar, Erman Rahman, menyebut hujan lebat tak hanya memicu banjir, tapi juga longsor di sejumlah titik, satu di antaranya di Pessel, Bukit Buah Palo.

"Hingga saat ini kami masih mendata korban banjir, tanah longsor dan jembatan putus," ujar Rahman, Jumat (12/10), kepada wartawan di Padang, Agusmanto, yang melapor untuk BBC News Indonesia.

Banjir yang terjadi membuat sebuah jembatan sepanjang delapan meter menuju Benteng Gunung Tak Jadi amblas.

Jembatan itu berada di Kecamatan Bonjol, Kabupaten Pasaman.

Camat Bonjol, Yohanis, menyebut jembatan tersebut tidak bisa dipakai lagi.

Ia berkata, banjir juga membuat jalan di sekitar jembatan rusak.

"Jembatan ini salah satu akses bagi pengendara sepeda motor dari arah pasar Bonjol menuju benteng dan perkapungan warga," kata Yohanis.

Yohanis menyebut jembatan itu berpotensi hanyut seiring banjir yang tak kunjung surut.

"Kami mengimbau warga Bonjol dan sekitarnya tidak melewati jembatan tersebut," ucapnya.

Jalan lintas Sumatera di Kelurahan Berangin, Bonjol, juga dilaporkan sempat terputus akibat longsor dan pohon kayu yang bertumbangan ke jalan.

Polisi kini menerapkan sistem buka-tutup di jalan itu.

Kepala BPBD Pasaman, Masfet Kenedi menuturkan, material longsor yang menimbun jalan di Berangin mencapai lebih dari setengah meter.

Pohon kayu sepanjang 18 meter juga tumbang ke jalan.

Sebuah mobil ambulans dilaporkan terjebak material longsor di Berangin, di jalur yang menghubungkan Bukittinggi dengan Sumatera Utara.

Mobil itu dalam posisi menanjak, lantas terjebak material longsor dari arah kiri.

Kepala Dinas Kesehatan Pasaman, Amdarisman, mengatakan, semua penumpang di mobil tersebut selamat, termasuk ibu hamil yang ada di dalam ambulans tersebut.

"Tidak ada korban jiwa.

Pasien (ibu hamil) sudah dirujuk ke RSUD Pasaman," ujarnya.

Longsor yang terjadi Ranah Batu, Jorong Piubuah, Nagari Tanjung Bonai, Kecamatan Lintau Buo Utara, Tanah Datar terjadi Kamis kemarin, sekitar pukul 18.30 WIB.

Dua orang dilaporkan tewas tertimbun material longsor, sementara dua orang lainnya lainnya hingga kini masih dalam pencarian.

Wali Nagari Jorong Piubuah, Haris, menyebut longsor menghancurkan setengah kampungnya.

Harie berkata, longsor terjadi secara cepat sehingga para sejumlah warga tak sempat menyelamatkan diri.

Longsor berasal dari Gunung Sago yang kemudian dengan cepat menyapu perkampungan di lembahnya.

"Ada dua korban, salah satunya ditemukan 2 kilometer dari rumahnya.

Korban ini sedang hamil," kata Haris.

Kerusakan di lembah Gunung Sago itu mencapai panjang dua kilometer dan sementara kedalaman lumpur berkisar 1,5 meter.

"Total korban terdampak sebanyak 26 jiwa, rata-rata terluka dan sudah dievakuasi ke rumah sakit," ujar Haris.

B.2 Hasil ringkasan artikel berita untuk evaluasi terbaik dengan *compression rate* 5%

Artikel asli berita 8:

Data Pusat Vulkanologi dan Mitigasi Bencana Geologi (PVMBG) mencatat, 80,1% dari 558 kecamatan di Jawa Barat tergolong rawan bencana tanah longsor.

Longsor diprediksi bisa menelan banyak korban karena pemerintah daerah dinilai gagal mengelola lereng yang makin sesak pemukiman.

"Jawa Barat penduduknya lebih dari 40 juta jiwa, mereka tinggal di perbukitan, dan hingga saat ini karena kebutuhan pemukiman, kami kutip pemerintah lalai," ujar Kepala Mitigasi Gerakan Tanah PVMBG, Agus Budianto, Rabu (7/2).

Kepada BBC Indonesia, Agus menyebut aktivitas manusia berpengaruh besar pada longsor di kawasan lereng, lembah sungai, dan dataran tinggi lainnya.

Agus berkata, aktivitas manusia itulah yang membedakan dampak bencana longsor di Jawa Barat dan daerah lain yang memiliki tanah gembur atau lapuk serta mudah menyerap air.

"Masyarakat mau tidak mau terus tumbuh di wilayah ini.

Lereng tidak stabil karena untuk perumahan," kata Agus.

Di luar faktor manusia, secara geologi, kata Agus, Jawa Barat berada di atas tanah yang mudah menyerap air dan gembur.

Ia menuturkan, faktor itu berkaitan pula dengan kondisi geografis Jabar.

"Ini daerah pegunungan, daerah gunung api, kalau melapuk sangat mudah pecah.

Seumpamanya hujan, tanah akan menyerap air sehingga licin dan gembur," kata Agus.

Gubernur Jawa Barat Ahmad Heryawan alias Aher tidak memungkiri data PVMBG.

Usai longsor di Kecamatan Cijeruk, Bogor, yang setidaknya menyebabkan lima warga meninggal dunia awal pekan ini, Aher mengklaim pihaknya sudah dan akan terus menormalisasi hutan di kawasan Puncak.

"Kami setiap tahun mencegah dengan cara menanam pohon, dengan biopori, sumber resapan, untuk menghijaukan kawasan-kawasan gundul termasuk daerah aliran Sungai Ciliwung," kata Aher kepada pers di Jakarta, Rabu siang (07/02).

Dalam catatan PVMBG, wilayah Jabar yang paling rawan longsor adalah Sukabumi (44 kecamatan), Garut (42), Bogor (35), dan Cianjur (31).

Adapun, wilayah Jabar yang memiliki potensi longsor paling rendah adalah Bekasi (2 kecamatan), Indramayu (2) dan Cirebon (1).

Merujuk data PVMBG per Februari 2018, Pulau Jawa merupakan wilayah dengan tingkat kerentanan longsor paling tinggi.

Persentase potensi longsor pada empat dari enam provinsi di pulau ini lebih dari 50%.

Berada satu peringkat di bawah Jabar, 78,2% kecamatan di Jawa Tengah berpotensi longsor dalam kategori menengah-tinggi.

Berturut-turut di peringkat setelahnya adalah Daerah Istimewa Yogyakarta (61,5%), Jawa Timur (57,9%), DKI Jakarta (22,7%), dan Banten (20%).

Agus Budiarto mengatakan PVMBG menyediakan peta kerawanan tanah bergerak agar dampak longsor dapat dihindari.

Salah satu peringatan yang diberikan PVMBG adalah memindahkan drainase atau arus air dari dataran tinggi yang menuju ke pemukiman penduduk.

"Gerakan tanah tidak bisa 100% dicegah, tapi jatuhnya korban dapat diminimalisir.

"Vegetasi harus diselesaikan dan air tidak mengalir ke lereng, melainkan ke tempat landai.

Masyarakat harus melaporkan kalau air yang meluncur terus membesar," kata Agus.

Dalam data PVMBG, Papua dan Papua Barat tergolong provinsi dengan jumlah kecamatan terbanyak yang rentan longsor.

Di Papua, lima kabupaten, yaitu Jaya Wijaya, Lanny Jaya, Pegunungan Bintang, Tolikara, dan Yahukimo, masuk kategori rawan longsor tertinggi.

Daerah tadi serupa dengan Fak-Fak, Sorong, Pegunungan Arfak, dan Teluk Bintuni di Papua Barat.

Sementara di Sulawesi Selatan, tingkat kerawanan yang sama ada di Tana Toraja dan Toraja Utara yang juga berada di perbukitan.

Agus Budianto berkata, meski secara geologi daerah-daerah itu rentan longsor, terutama saat curah hujan sedang tinggi, aktivitas manusia yang minim dapat meminimalisir dampak bencana alam.

"Papua memang ada perbedaan dari tata guna lahan, tapi kalau komponen geologi--perbukitan, material, dan drainase--potensi longsohnya sama," tuturnya.


Sejak awal tahun 2018, sejumlah bencana longsor telah terjadi di berbagai daerah.


Januari lalu, longsor dilaporkan melanda Banjarnegara, Wonogiri, hingga Magelang.

Sementara sejak akhir pekan lalu, longsor menerjang Kudus, Jepara, dan Bogor.

Kalimat yang dipilih sebagai ringkasan pada artikel berita 8:

Keterangan :

 : Kalimat terpilih sebagai ringkasan oleh pakar

 : Kalimat terpilih sebagai ringkasan oleh sistem

 : Kalimat terpilih sebagai ringkasan oleh sistem dan pakar

Data Pusat Vulkanologi dan Mitigasi Bencana Geologi (PVMBG) mencatat, 80,1% dari 558 kecamatan di Jawa Barat tergolong rawan bencana tanah longsor.

Longsor diprediksi bisa menelan banyak korban karena pemerintah daerah dinilai gagal mengelola lereng yang makin sesak pemukiman.

"Jawa Barat penduduknya lebih dari 40 juta jiwa, mereka tinggal di perbukitan, dan hingga saat ini karena kebutuhan pemukiman, kami kutip pemerintah lalai," ujar Kepala Mitigasi Gerakan Tanah PVMBG, Agus Budianto, Rabu (7/2).

Kepada BBC Indonesia, Agus menyebut aktivitas manusia berpengaruh besar pada longsor di kawasan lereng, lembah sungai, dan dataran tinggi lainnya.

Agus berkata, aktivitas manusia itulah yang membedakan dampak bencana longsor di Jawa Barat dan daerah lain yang memiliki tanah gembur atau lapuk serta mudah menyerap air.

"Masyarakat mau tidak mau terus tumbuh di wilayah ini.

Lereng tidak stabil karena untuk perumahan," kata Agus.

Di luar faktor manusia, secara geologi, kata Agus, Jawa Barat berada di atas tanah yang mudah menyerap air dan gembur.

Ia menuturkan, faktor itu berkaitan pula dengan kondisi geografis Jabar.

"Ini daerah pegunungan, daerah gunung api, kalau melapuk sangat mudah pecah.

Seumpamanya hujan, tanah akan menyerap air sehingga licin dan gembur," kata Agus.

Gubernur Jawa Barat Ahmad Heryawan alias Aher tidak memungkiri data PVMBG.

Usai longsor di Kecamatan Cijeruk, Bogor, yang setidaknya menyebabkan lima warga meninggal dunia awal pekan ini, Aher mengklaim pihaknya sudah dan akan terus menormalisasi hutan di kawasan Puncak.

"Kami setiap tahun mencegah dengan cara menanam pohon, dengan biopori, sumber resapan, untuk menghijaukan kawasan-kawasan gundul termasuk daerah aliran Sungai Ciliwung," kata Aher kepada pers di Jakarta, Rabu siang (07/02).

Dalam catatan PVMBG, wilayah Jabar yang paling rawan longsor adalah Sukabumi (44 kecamatan), Garut (42), Bogor (35), dan Cianjur (31).

Adapun, wilayah Jabar yang memiliki potensi longsor paling rendah adalah Bekasi (2 kecamatan), Indramayu (2) dan Cirebon (1).

Merujuk data PVMBG per Februari 2018, Pulau Jawa merupakan wilayah dengan tingkat kerentanan longsor paling tinggi.

Persentase potensi longsor pada empat dari enam provinsi di pulau ini lebih dari 50%.

Berada satu peringkat di bawah Jabar, 78,2% kecamatan di Jawa Tengah berpotensi longsor dalam kategori menengah-tinggi.

Berturut-turut di peringkat setelahnya adalah Daerah Istimewa Yogyakarta (61,5%), Jawa Timur (57,9%), DKI Jakarta (22,7%), dan Banten (20%).

Agus Budiarto mengatakan PVMBG menyediakan peta kerawanan tanah bergerak agar dampak longsor dapat dihindari.

Salah satu peringatan yang diberikan PVMBG adalah memindahkan drainase atau arus air dari dataran tinggi yang menuju ke pemukiman penduduk.

"Gerakan tanah tidak bisa 100% dicegah, tapi jatuhnya korban dapat diminimalisir.

"Vegetasi harus diselesaikan dan air tidak mengalir ke lereng, melainkan ke tempat landai.

Masyarakat harus melaporkan kalau air yang meluncur terus membesar," kata Agus.

Dalam data PVMBG, Papua dan Papua Barat tergolong provinsi dengan jumlah kecamatan terbanyak yang rentan longsor.

Di Papua, lima kabupaten, yaitu Jaya Wijaya, Lanny Jaya, Pegunungan Bintang, Tolikara, dan Yahukimo, masuk kategori rawan longsor tertinggi.

Daerah tadi serupa dengan Fak-Fak, Sorong, Pegunungan Arfak, dan Teluk Bintuni di Papua Barat.

Sementara di Sulawesi Selatan, tingkat kerawanan yang sama ada di Tana Toraja dan Toraja Utara yang juga berada di perbukitan.

Agus Budianto berkata, meski secara geologi daerah-daerah itu rentan longsor, terutama saat curah hujan sedang tinggi, aktivitas manusia yang minim dapat meminimalisir dampak bencana alam.

"Papua memang ada perbedaan dari tata guna lahan, tapi kalau komponen geologi--perbukitan, material, dan drainase--potensi longsohnya sama," tuturnya.

Sejak awal tahun 2018, sejumlah bencana longsor telah terjadi di berbagai daerah.

Januari lalu, longsor dilaporkan melanda Banjarnegara, Wonogiri, hingga Magelang.

Sementara sejak akhir pekan lalu, longsor menerjang Kudus, Jepara, dan Bogor.

B.3 Hasil ringkasan artikel berita untuk evaluasi terburuk dengan *compression rate 10%*

Artikel asli berita 4:

Sejumlah kabupaten dan kota di Sumatera Barat dilanda banjir seiring hujan lebat yang terjadi sejak Kamis (11/10).

Bencana alam ini dilaporkan menyebabkan enam orang tewas dan ratusan rumah terendam air dengan kedalaman hingga satu meter.

Merujuk data Pusat Pengendalian Operasi Penanggulangan Bencana (Pusdalops) BPBD Sumbar, tiga korban tewas berada di Padang Pariaman, sementara tiga lainnya di Tanah Datar.

Adapun, daerah terdampak banjir adalah Tanah Datar, Limapuluh Kota, Pesisir Selatan, Pasaman, Pasaman Barat, Padang Pariaman, Agam, Mentawai Sijunjung, Solok dan Sawahlunto.

Kepala BPBD Sumbar, Erman Rahman, menyebut hujan lebat tak hanya memicu banjir, tapi juga longsor di sejumlah titik, satu di antaranya di Pessel, Bukit Buah Palo.

"Hingga saat ini kami masih mendata korban banjir, tanah longsor dan jembatan putus," ujar Rahman, Jumat (12/10), kepada wartawan di Padang, Agusmanto, yang melapor untuk BBC News Indonesia.

Banjir yang terjadi membuat sebuah jembatan sepanjang delapan meter menuju Benteng Gunung Tak Jadi ambles.

Jembatan itu berada di Kecamatan Bonjol, Kabupaten Pasaman.

Camat Bonjol, Yohanis, menyebut jembatan tersebut tidak bisa dipakai lagi.

la berkata, banjir juga membuat jalan di sekitar jembatan rusak.

"Jembatan ini salah satu akses bagi pengendara sepeda motor dari arah pasar Bonjol menuju benteng dan perkapungan warga," kata Yohanis.

Yohanis menyebut jembatan itu berpotensi hanyut seiring banjir yang tak kunjung surut.

"Kami mengimbau warga Bonjol dan sekitarnya tidak melewati jembatan tersebut," ucapnya.

Jalan lintas Sumatera di Kelurahan Berangin, Bonjol, juga dilaporkan sempat terputus akibat longsor dan pohon kayu yang bertumbangan ke jalan.

Polisi kini menerapkan sistem buka-tutup di jalan itu.

Kepala BPBD Pasaman, Masfet Kenedi menuturkan, material longsor yang menimbun jalan di Berangin mencapai lebih dari setengah meter.

Pohon kayu sepanjang 18 meter juga tumbang ke jalan.

Sebuah mobil ambulans dilaporkan terjebak material longsor di Berangin, di jalur yang menghubungkan Bukittinggi dengan Sumatera Utara.

Mobil itu dalam posisi menanjak, lantas terjebak material longsor dari arah kiri.

Kepala Dinas Kesehatan Pasaman, Amdarisman, mengatakan, semua penumpang di mobil tersebut selamat, termasuk ibu hamil yang ada di dalam ambulans tersebut.

"Tidak ada korban jiwa.

Pasien (ibu hamil) sudah dirujuk ke RSUD Pasaman," ujarinya.

Longsor yang terjadi Ranah Batu, Jorong Piubuah, Nagari Tanjung Bonai, Kecamatan Lintau Buo Utara, Tanah Datar terjadi Kamis kemarin, sekitar pukul 18.30 WIB.

Dua orang dilaporkan tewas tertimbun material longsor, sementara dua orang lainnya lainnya hingga kini masih dalam pencarian.

Wali Nagari Jorong Piubuah, Haris, menyebut longsor menghancurkan setengah kampungnya.

Harie berkata, longsor terjadi secara cepat sehingga para sejumlah warga tak sempat menyelamatkan diri.

Longsor berasal dari Gunung Sago yang kemudian dengan cepat menyapu perkampungan di lembahnya.

"Ada dua korban, salah satunya ditemukan 2 kilometer dari rumahnya.

Korban ini sedang hamil," kata Haris.

Kerusakan di lembah Gunung Sago itu mencapai panjang dua kilometer dan sementara kedalaman lumpur berkisar 1,5 meter.

"Total korban terdampak sebanyak 26 jiwa, rata-rata terluka dan sudah dievakuasi ke rumah sakit," ujar Haris.

Kalimat yang dipilih sebagai ringkasan pada artikel berita 4:

Keterangan :

 : Kalimat terpilih sebagai ringkasan oleh pakar

 : Kalimat terpilih sebagai ringkasan oleh sistem

Sejumlah kabupaten dan kota di Sumatera Barat dilanda banjir seiring hujan lebat yang terjadi sejak Kamis (11/10).

Bencana alam ini dilaporkan menyebabkan enam orang tewas dan ratusan rumah terendam air dengan kedalaman hingga satu meter.

Merujuk data Pusat Pengendalian Operasi Penanggulangan Bencana (Pusdalops) BPBD Sumbar, tiga korban tewas berada di Padang Pariaman, sementara tiga lainnya di Tanah Datar.

Adapun, daerah terdampak banjir adalah Tanah Datar, Limapuluh Kota, Pesisir Selatan, Pasaman, Pasaman Barat, Padang Pariaman, Agam, Mentawai Sijunjung, Solok dan Sawahlunto.

Kepala BPBD Sumbar, Erman Rahman, menyebut hujan lebat tak hanya memicu banjir, tapi juga longsor di sejumlah titik, satu di antaranya di Pessel, Bukit Buah Palo.

"Hingga saat ini kami masih mendata korban banjir, tanah longsor dan jembatan putus," ujar Rahman, Jumat (12/10), kepada wartawan di Padang, Agusmanto, yang melapor untuk BBC News Indonesia.

Banjir yang terjadi membuat sebuah jembatan sepanjang delapan meter menuju Benteng Gunung Tak Jadi amblas.

Jembatan itu berada di Kecamatan Bonjol, Kabupaten Pasaman.

Camat Bonjol, Yohanis, menyebut jembatan tersebut tidak bisa dipakai lagi.

Ia berkata, banjir juga membuat jalan di sekitar jembatan rusak.

"Jembatan ini salah satu akses bagi pengendara sepeda motor dari arah pasar Bonjol menuju benteng dan perkapungan warga," kata Yohanis.

Yohanis menyebut jembatan itu berpotensi hanyut seiring banjir yang tak kunjung surut.

"Kami mengimbau warga Bonjol dan sekitarnya tidak melewati jembatan tersebut," ucapnya.

Jalan lintas Sumatera di Kelurahan Berangin, Bonjol, juga dilaporkan sempat terputus akibat longsor dan pohon kayu yang bertumbangan ke jalan.

Polisi kini menerapkan sistem buka-tutup di jalan itu.

Kepala BPBD Pasaman, Masfet Kenedi menuturkan, material longsor yang menimbun jalan di Berangin mencapai lebih dari setengah meter.

Pohon kayu sepanjang 18 meter juga tumbang ke jalan.

Sebuah mobil ambulans dilaporkan terjebak material longsor di Berangin, di jalur yang menghubungkan Bukittinggi dengan Sumatera Utara.

Mobil itu dalam posisi menanjak, lantas terjebak material longsor dari arah kiri.

Kepala Dinas Kesehatan Pasaman, Amdarisman, mengatakan, semua penumpang di mobil tersebut selamat, termasuk ibu hamil yang ada di dalam ambulans tersebut.

"Tidak ada korban jiwa.

Pasien (ibu hamil) sudah dirujuk ke RSUD Pasaman," ujarnya.

Longsor yang terjadi Ranah Batu, Jorong Piubuah, Nagari Tanjung Bonai, Kecamatan Lintau Buo Utara, Tanah Datar terjadi Kamis kemarin, sekitar pukul 18.30 WIB.

Dua orang dilaporkan tewas tertimbun material longsor, sementara dua orang lainnya lainnya hingga kini masih dalam pencarian.

Wali Nagari Jorong Piubuah, Haris, menyebut longsor menghancurkan setengah kampungnya.

Harie berkata, longsor terjadi secara cepat sehingga para sejumlah warga tak sempat menyelamatkan diri.

Longsor berasal dari Gunung Sago yang kemudian dengan cepat menyapu perkampungan di lembahnya.

"Ada dua korban, salah satunya ditemukan 2 kilometer dari rumahnya.

Korban ini sedang hamil," kata Haris.

Kerusakan di lembah Gunung Sago itu mencapai panjang dua kilometer dan sementara kedalaman lumpur berkisar 1,5 meter.

"Total korban terdampak sebanyak 26 jiwa, rata-rata terluka dan sudah dievakuasi ke rumah sakit," ujar Haris.

B.4 Hasil ringkasan artikel berita untuk evaluasi terbaik dengan *compression rate 10%*

Artikel asli berita 8:

Data Pusat Vulkanologi dan Mitigasi Bencana Geologi (PVMBG) mencatat, 80,1% dari 558 kecamatan di Jawa Barat tergolong rawan bencana tanah longsor.

Longsor diprediksi bisa menelan banyak korban karena pemerintah daerah dinilai gagal mengelola lereng yang makin sesak pemukiman.

"Jawa Barat penduduknya lebih dari 40 juta jiwa, mereka tinggal di perbukitan, dan hingga saat ini karena kebutuhan pemukiman, kami kutip pemerintah lalai," ujar Kepala Mitigasi Gerakan Tanah PVMBG, Agus Budiarto, Rabu (7/2).

Kepada BBC Indonesia, Agus menyebut aktivitas manusia berpengaruh besar pada longsor di kawasan lereng, lembah sungai, dan dataran tinggi lainnya.

Agus berkata, aktivitas manusia itulah yang membedakan dampak bencana longsor di Jawa Barat dan daerah lain yang memiliki tanah gembur atau lapuk serta mudah menyerap air.

"Masyarakat mau tidak mau terus tumbuh di wilayah ini.

Lereng tidak stabil karena untuk perumahan," kata Agus.

Di luar faktor manusia, secara geologi, kata Agus, Jawa Barat berada di atas tanah yang mudah menyerap air dan gembur.

Ia menuturkan, faktor itu berkaitan pula dengan kondisi geografis Jabar.

"Ini daerah pegunungan, daerah gunung api, kalau melapuk sangat mudah pecah.

Seumpamanya hujan, tanah akan menyerap air sehingga licin dan gembur," kata Agus.

Gubernur Jawa Barat Ahmad Heryawan alias Aher tidak memungkiri data PVMBG.

Usai longsor di Kecamatan Cijeruk, Bogor, yang setidaknya menyebabkan lima warga meninggal dunia awal pekan ini, Aher mengklaim pihaknya sudah dan akan terus menormalisasi hutan di kawasan Puncak.

"Kami setiap tahun mencegah dengan cara menanam pohon, dengan biopori, sumber resapan, untuk menghijaukan kawasan-kawasan gundul termasuk daerah aliran Sungai Ciliwung," kata Aher kepada pers di Jakarta, Rabu siang (07/02).

Dalam catatan PVMBG, wilayah Jabar yang paling rawan longsor adalah Sukabumi (44 kecamatan), Garut (42), Bogor (35), dan Cianjur (31).

Adapun, wilayah Jabar yang memiliki potensi longsor paling rendah adalah Bekasi (2 kecamatan), Indramayu (2) dan Cirebon (1).

Merujuk data PVMBG per Februari 2018, Pulau Jawa merupakan wilayah dengan tingkat kerentanan longsor paling tinggi.

Persentase potensi longsor pada empat dari enam provinsi di pulau ini lebih dari 50%.

Berada satu peringkat di bawah Jabar, 78,2% kecamatan di Jawa Tengah berpotensi longsor dalam kategori menengah-tinggi.

Berturut-turut di peringkat setelahnya adalah Daerah Istimewa Yogyakarta (61,5%), Jawa Timur (57,9%), DKI Jakarta (22,7%), dan Banten (20%).

Agus Budiarto mengatakan PVMBG menyediakan peta kerawanan tanah bergerak agar dampak longsor dapat dihindari.

Salah satu peringatan yang diberikan PVMBG adalah memindahkan drainase atau arus air dari dataran tinggi yang menuju ke pemukiman penduduk.

"Gerakan tanah tidak bisa 100% dicegah, tapi jatuhnya korban dapat diminimalisir.

"Vegetasi harus diselesaikan dan air tidak mengalir ke lereng, melainkan ke tempat landai.

Masyarakat harus melaporkan kalau air yang meluncur terus membesar," kata Agus.

Dalam data PVMBG, Papua dan Papua Barat tergolong provinsi dengan jumlah kecamatan terbanyak yang rentan longsor.

Di Papua, lima kabupaten, yaitu Jaya Wijaya, Lanny Jaya, Pegunungan Bintang, Tolikara, dan Yahukimo, masuk kategori rawan longsor tertinggi.

Daerah tadi serupa dengan Fak-Fak, Sorong, Pegunungan Arfak, dan Teluk Bintuni di Papua Barat.

Sementara di Sulawesi Selatan, tingkat kerawanan yang sama ada di Tana Toraja dan Toraja Utara yang juga berada di perbukitan.

Agus Budianto berkata, meski secara geologi daerah-daerah itu rentan longsor, terutama saat curah hujan sedang tinggi, aktivitas manusia yang minim dapat meminimalisir dampak bencana alam.

"Papua memang ada perbedaan dari tata guna lahan, tapi kalau komponen geologi--perbukitan, material, dan drainase--potensi longsohnya sama," tuturnya.


Sejak awal tahun 2018, sejumlah bencana longsor telah terjadi di berbagai daerah.


Januari lalu, longsor dilaporkan melanda Banjarnegara, Wonogiri, hingga Magelang.

Sementara sejak akhir pekan lalu, longsor menerjang Kudus, Jepara, dan Bogor.

Kalimat yang dipilih sebagai ringkasan pada artikel berita 8:

Keterangan :

 : Kalimat terpilih sebagai ringkasan oleh pakar

 : Kalimat terpilih sebagai ringkasan oleh sistem

 : Kalimat terpilih sebagai ringkasan oleh sistem dan pakar

Data Pusat Vulkanologi dan Mitigasi Bencana Geologi (PVMBG) mencatat, 80,1% dari 558 kecamatan di Jawa Barat tergolong rawan bencana tanah longsor.

Longsor diprediksi bisa menelan banyak korban karena pemerintah daerah dinilai gagal mengelola lereng yang makin sesak pemukiman.

"Jawa Barat penduduknya lebih dari 40 juta jiwa, mereka tinggal di perbukitan, dan hingga saat ini karena kebutuhan pemukiman, kami kutip pemerintah lalai," ujar Kepala Mitigasi Gerakan Tanah PVMBG, Agus Budianto, Rabu (7/2).

Kepada BBC Indonesia, Agus menyebut aktivitas manusia berpengaruh besar pada longsor di kawasan lereng, lembah sungai, dan dataran tinggi lainnya.

Agus berkata, aktivitas manusia itulah yang membedakan dampak bencana longsor di Jawa Barat dan daerah lain yang memiliki tanah gembur atau lapuk serta mudah menyerap air.

"Masyarakat mau tidak mau terus tumbuh di wilayah ini. Lereng tidak stabil karena untuk perumahan," kata Agus.

Di luar faktor manusia, secara geologi, kata Agus, Jawa Barat berada di atas tanah yang mudah menyerap air dan gembur.

Ia menuturkan, faktor itu berkaitan pula dengan kondisi geografis Jabar.

"Ini daerah pegunungan, daerah gunung api, kalau melapuk sangat mudah pecah.

Seumpamanya hujan, tanah akan menyerap air sehingga licin dan gembur," kata Agus.

Gubernur Jawa Barat Ahmad Heryawan alias Aher tidak memungkiri data PVMBG.

Usai longsor di Kecamatan Cijeruk, Bogor, yang setidaknya menyebabkan lima warga meninggal dunia awal pekan ini, Aher mengklaim pihaknya sudah dan akan terus menormalisasi hutan di kawasan Puncak.

"Kami setiap tahun mencegah dengan cara menanam pohon, dengan biopori, sumber resapan, untuk menghijaukan kawasan-kawasan gundul termasuk daerah aliran Sungai Ciliwung," kata Aher kepada pers di Jakarta, Rabu siang (07/02).

Dalam catatan PVMBG, wilayah Jabar yang paling rawan longsor adalah Sukabumi (44 kecamatan), Garut (42), Bogor (35), dan Cianjur (31).

Adapun, wilayah Jabar yang memiliki potensi longsor paling rendah adalah Bekasi (2 kecamatan), Indramayu (2) dan Cirebon (1).

Merujuk data PVMBG per Februari 2018, Pulau Jawa merupakan wilayah dengan tingkat kerentanan longsor paling tinggi.

Persentase potensi longsor pada empat dari enam provinsi di pulau ini lebih dari 50%.

Berada satu peringkat di bawah Jabar, 78,2% kecamatan di Jawa Tengah berpotensi longsor dalam kategori menengah-tinggi.

Berturut-turut di peringkat setelahnya adalah Daerah Istimewa Yogyakarta (61,5%), Jawa Timur (57,9%), DKI Jakarta (22,7%), dan Banten (20%).

Agus Budianto mengatakan PVMBG menyediakan peta kerawanan tanah bergerak agar dampak longsor dapat dinihalkan.

Salah satu peringatan yang diberikan PVMBG adalah memindahkan drainase atau arus air dari dataran tinggi yang menuju ke pemukiman penduduk.

"Gerakan tanah tidak bisa 100% dicegah, tapi jatuhnya korban dapat diminimalisir.

"Vegetasi harus diselesaikan dan air tidak mengalir ke lereng, melainkan ke tempat landai.

Masyarakat harus melaporkan kalau air yang meluncur terus membesar," kata Agus.

Dalam data PVMBG, Papua dan Papua Barat tergolong provinsi dengan jumlah kecamatan terbanyak yang rentan longsor.

Di Papua, lima kabupaten, yaitu Jaya Wijaya, Lanny Jaya, Pegunungan Bintang, Tolikara, dan Yahukimo, masuk kategori rawan longsor tertinggi.

Daerah tadi serupa dengan Fak-Fak, Sorong, Pegunungan Arfak, dan Teluk Bintuni di Papua Barat.

Sementara di Sulawesi Selatan, tingkat kerawanan yang sama ada di Tana Toraja dan Toraja Utara yang juga berada di perbukitan.

Agus Budianto berkata, meski secara geologi daerah-daerah itu rentan longsor, terutama saat curah hujan sedang tinggi, aktivitas manusia yang minim dapat meminimalisir dampak bencana alam.

"Papua memang ada perbedaan dari tata guna lahan, tapi kalau komponen geologi--perbukitan, material, dan drainase--potensi longsornya sama," tuturnya.

Sejak awal tahun 2018, sejumlah bencana longsor telah terjadi di berbagai daerah.

Januari lalu, longsor dilaporkan melanda Banjarnegara, Wonogiri, hingga Magelang.

Sementara sejak akhir pekan lalu, longsor menerjang Kudus, Jepara, dan Bogor.

B.5 Hasil ringkasan artikel berita untuk evaluasi terburuk dengan *compression rate 20%*

Artikel asli berita 4:

Sejumlah kabupaten dan kota di Sumatera Barat dilanda banjir seiring hujan lebat yang terjadi sejak Kamis (11/10).

Bencana alam ini dilaporkan menyebabkan enam orang tewas dan ratusan rumah terendam air dengan kedalaman hingga satu meter.

Merujuk data Pusat Pengendalian Operasi Penanggulangan Bencana (Pusdalops) BPBD Sumbar, tiga korban tewas berada di Padang Pariaman, sementara tiga lainnya di Tanah Datar.

Adapun, daerah terdampak banjir adalah Tanah Datar, Limapuluh Kota, Pesisir Selatan, Pasaman, Pasaman Barat, Padang Pariaman, Agam, Mentawai Sijunjung, Solok dan Sawahlunto.

Kepala BPBD Sumbar, Erman Rahman, menyebut hujan lebat tak hanya memicu banjir, tapi juga longsor di sejumlah titik, satu di antaranya di Pessel, Bukit Buah Palo.

"Hingga saat ini kami masih mendata korban banjir, tanah longsor dan jembatan putus," ujar Rahman, Jumat (12/10), kepada wartawan di Padang, Agusmanto, yang melapor untuk BBC News Indonesia.

Banjir yang terjadi membuat sebuah jembatan sepanjang delapan meter menuju Benteng Gunung Tak Jadi amblas.

Jembatan itu berada di Kecamatan Bonjol, Kabupaten Pasaman.

Camat Bonjol, Yohanis, menyebut jembatan tersebut tidak bisa dipakai lagi.

Ia berkata, banjir juga membuat jalan di sekitar jembatan rusak.

"Jembatan ini salah satu akses bagi pengendara sepeda motor dari arah pasar Bonjol menuju benteng dan perkapungan warga," kata Yohanis.

Yohanis menyebut jembatan itu berpotensi hanyut seiring banjir yang tak kunjung surut.

"Kami mengimbau warga Bonjol dan sekitarnya tidak melewati jembatan tersebut," ucapnya.

Jalan lintas Sumatera di Kelurahan Berangin, Bonjol, juga dilaporkan sempat terputus akibat longsor dan pohon kayu yang bertumbangan ke jalan.

Polisi kini menerapkan sistem buka-tutup di jalan itu.

Kepala BPBD Pasaman, Masfet Kenedi menuturkan, material longsor yang menimbun jalan di Berangin mencapai lebih dari setengah meter.

Pohon kayu sepanjang 18 meter juga tumbang ke jalan.

Sebuah mobil ambulans dilaporkan terjebak material longsor di Berangin, di jalur yang menghubungkan Bukittinggi dengan Sumatera Utara.

Mobil itu dalam posisi menanjak, lantas terjebak material longsor dari arah kiri.

Kepala Dinas Kesehatan Pasaman, Amdarisman, mengatakan, semua penumpang di mobil tersebut selamat, termasuk ibu hamil yang ada di dalam ambulans tersebut.

"Tidak ada korban jiwa.

Pasien (ibu hamil) sudah dirujuk ke RSUD Pasaman," ujarnya.

Longsor yang terjadi Ranah Batu, Jorong Piubuah, Nagari Tanjung Bonai, Kecamatan Lintau Buo Utara, Tanah Datar terjadi Kamis kemarin, sekitar pukul 18.30 WIB.

Dua orang dilaporkan tewas tertimbun material longsor, sementara dua orang lainnya lainnya hingga kini masih dalam pencarian.

Wali Nagari Jorong Piubuah, Haris, menyebut longsor menghancurkan setengah kampungnya.

Harie berkata, longsor terjadi secara cepat sehingga para sejumlah warga tak sempat menyelamatkan diri.

Longsor berasal dari Gunung Sago yang kemudian dengan cepat menyapu perkampungan di lembahnya.

"Ada dua korban, salah satunya ditemukan 2 kilometer dari rumahnya.


Korban ini sedang hamil," kata Haris.


Kerusakan di lembah Gunung Sago itu mencapai panjang dua kilometer dan sementara kedalaman lumpur berkisar 1,5 meter.

"Total korban terdampak sebanyak 26 jiwa, rata-rata terluka dan sudah dievakuasi ke rumah sakit," ujar Haris.

Kalimat yang dipilih sebagai ringkasan pada artikel berita 4:

Keterangan :

 : Kalimat terpilih sebagai ringkasan oleh pakar

 : Kalimat terpilih sebagai ringkasan oleh sistem

 : Kalimat terpilih sebagai ringkasan oleh sistem dan pakar

Sejumlah kabupaten dan kota di Sumatera Barat dilanda banjir seiring hujan lebat yang terjadi sejak Kamis (11/10).

Bencana alam ini dilaporkan menyebabkan enam orang tewas dan ratusan rumah terendam air dengan kedalaman hingga satu meter.

Merujuk data Pusat Pengendalian Operasi Penanggulangan Bencana (Pusdalops) BPBD Sumbar, tiga korban tewas berada di Padang Pariaman, sementara tiga lainnya di Tanah Datar.

Adapun, daerah terdampak banjir adalah Tanah Datar, Limapuluh Kota, Pesisir Selatan, Pasaman, Pasaman Barat, Padang Pariaman, Agam, Mentawai Sijunjung, Solok dan Sawahlunto.

Kepala BPBD Sumbar, Erman Rahman, menyebut hujan lebat tak hanya memicu banjir, tapi juga longsor di sejumlah titik, satu di antaranya di Pessel, Bukit Buah Palo.

"Hingga saat ini kami masih mendata korban banjir, tanah longsor dan jembatan putus," ujar Rahman, Jumat (12/10), kepada wartawan di Padang, Agusmanto, yang melapor untuk BBC News Indonesia.

Banjir yang terjadi membuat sebuah jembatan sepanjang delapan meter menuju Benteng Gunung Tak Jadi amblas.

Jembatan itu berada di Kecamatan Bonjol, Kabupaten Pasaman.

Camat Bonjol, Yohanis, menyebut jembatan tersebut tidak bisa dipakai lagi.

Ia berkata, banjir juga membuat jalan di sekitar jembatan rusak.

"Jembatan ini salah satu akses bagi pengendara sepeda motor dari arah pasar Bonjol menuju benteng dan perkapungan warga," kata Yohanis.

Yohanis menyebut jembatan itu berpotensi hanyut seiring banjir yang tak kunjung surut.

"Kami mengimbau warga Bonjol dan sekitarnya tidak melewati jembatan tersebut," ucapnya.

Jalan lintas Sumatera di Kelurahan Berangin, Bonjol, juga dilaporkan sempat terputus akibat longsor dan pohon kayu yang bertumbangan ke jalan.

Polisi kini menerapkan sistem buka-tutup di jalan itu.

Kepala BPBD Pasaman, Masfet Kenedi menuturkan, material longsor yang menimbun jalan di Berangin mencapai lebih dari setengah meter.

Pohon kayu sepanjang 18 meter juga tumbang ke jalan.

Sebuah mobil ambulans dilaporkan terjebak material longsor di Berangin, di jalur yang menghubungkan Bukittinggi dengan Sumatera Utara.

Mobil itu dalam posisi menanjak, lantas terjebak material longsor dari arah kiri.

Kepala Dinas Kesehatan Pasaman, Amdarisman, mengatakan, semua penumpang di mobil tersebut selamat, termasuk ibu hamil yang ada di dalam ambulans tersebut.

"Tidak ada korban jiwa.

Pasien (ibu hamil) sudah dirujuk ke RSUD Pasaman," ujarnya.

Longsor yang terjadi Ranah Batu, Jorong Piubuah, Nagari Tanjung Bonai, Kecamatan Lintau Buo Utara, Tanah Datar terjadi Kamis kemarin, sekitar pukul 18.30 WIB.

Dua orang dilaporkan tewas tertimbun material longsor, sementara dua orang lainnya lainnya hingga kini masih dalam pencarian.

Wali Nagari Jorong Piubuah, Haris, menyebut longsor menghancurkan setengah kampungnya.

Harie berkata, longsor terjadi secara cepat sehingga para sejumlah warga tak sempat menyelamatkan diri.

Longsor berasal dari Gunung Sago yang kemudian dengan cepat menyapu perkampungan di lembahnya.

"Ada dua korban, salah satunya ditemukan 2 kilometer dari rumahnya.

Korban ini sedang hamil," kata Haris.

Kerusakan di lembah Gunung Sago itu mencapai panjang dua kilometer dan sementara kedalaman lumpur berkisar 1,5 meter.

"Total korban terdampak sebanyak 26 jiwa, rata-rata terluka dan sudah dievakuasi ke rumah sakit," ujar Haris.

B.6 Hasil ringkasan artikel berita untuk evaluasi terbaik dengan *compression rate* 20%

Artikel asli berita 7:

Sejumlah sekolah di Kota Palembang, Sumatera Selatan, memutuskan memulangkan siswa-siswa mereka lantaran kabut asap semakin tebal menyelimuti kota tersebut pada Senin (14/10) pagi.

Kepala SMP Negeri 7 Palembang, Siti Zubaida, mengatakan keputusan pemulangan ditempuh sesuai dengan instruksi Dinas Pendidikan Kota Palembang.

"Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan berdampak buruk terhadap siswa, oleh karenanya atas instruksi Kadiknas Kota Palembang melalui pesan WA Grup meminta siswa dipulangkan dan belajar di rumah masing-masing saja," jelas Siti kepada radio Elshinta.

Hal ini diamini Kepala Dinas Pendidikan Kota Palembang, Ahmad Zulinto, yang menyampaikan surat edaran ke semua sekolah.

"Hari ini seluruh TK hingga SMP negeri dan swasta sederajat diliburkan, untuk besok dan seterusnya akan diberikan edaran lebih lanjut," kata Ahmad Zulinto kepada kantor berita Antara.

Menurutnya, kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan, kegiatan belajar mengajar di tingkat TK sampai SMP akan tetap diliburkan.

Akan tetapi, sebagaimana dipaparkan Kepala Dinas Pendidikan Sumatera Selatan, Widodo, kegiatan belajar mengajar di daerah yang tidak terdampak kabut asap tetap berlangsung.

"Daerah yang tidak terdampak kabut asap tetap normal tetap belajar, untuk daerah yang terkategori sedang tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker, bagi daerah terkategori parah maka siswa diberikan tugas dengan memaksimalkan kelas daring," kata Widodo kepada Antara.

Hal itu belakangan dibenarkan Agus Wibowo, selaku Kepala Pusat Data, Informasi dan Humas Badan Nasional Penanggulangan Bencana (BNPB).

"Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud, TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers.

Sejumlah warga Palembang, Sumatera Selatan, mengeluhkan kabut asap pada Senin (14/10) tergolong parah.

Bahkan jarak pandang hanya 10 meter.

Amelia, warga Kelurahan Bukit Sangkal, Kecamatan Kalidoni, mengatakan dirinya terkejut ketika akan ke luar rumah sekitar pukul 06.30 WIB untuk mengantar anaknya ke sekolah mendapati kabut sedemikian pekat.

"Saya terkejut, kenapa gelap ini.

Kemarin-kemarin ada kabut asap, tapi tidak separah hari ini," kata dia sebagaimana dikutip Antara.

Keluhan atas kondisi ini juga diungkapkan Tina, seorang guru senam di sebuah tempat kebugaran.

"Saya selalu ke luar rumah pukul 06.00 WIB karena ada jadwal senam, sempat terkejut juga karena jarak pandang hanya 10 meter.

Sangat terasa, apalagi saya pakai sepeda motor," kata dia.

Kepala Seksi Observasi dan Informasi Stasiun Meteorologi SMB II Palembang, Bambang Beny Setiaji, mengatakan kabut tersebut bercampur asap kiriman dari wilayah Kabupaten Ogan Komering Ilir (OKI) yang berada sebelah tenggara Kota Palembang.

Kabut asap di Kota Palembang semakin parah dalam sepekan terakhir akibat dampak kebakaran hutan dan lahan di sejumlah kabupaten.

Berdasarkan data Badan Penanggulangan Bencana Daerah Provinsi Sumsel yang bersumber dari Satelit Lapan disebutkan jumlah titik panas pada Senin (14/10) mencapai 732 titik, dengan titik panas terbanyak di Kabupaten Ogan Komering Ilir yang berjumlah 437 titik.

Sebelumnya, pada Jumat (11/10), titik panas berjumlah 417 titik.

Kepala Bidang Kedaruratan Badan Penanggulangan Bencana Daerah Provinsi Sumsel, Ansori, mengatakan titik panas terbanyak terpantau di Kabupaten Ogan Komering Ilir sehingga fokus pemadaman difokuskan di wilayah tersebut.

"Kami selalu lakukan waterbombing (pemadaman dari udara), setiap hari mengerahkan lima unit helikopter.

Kebakaran di OKI ini memang sulit dipadamkan karena terjadi di kawasan gambut, dan akses darat yang terbatas.

belum lagi jika terbakar, asapnya mengarah ke Palembang," kata Ansori.

Sementara itu, aktivitas kapal bertonase di Sungai Musi, Kota Palembang, dihentikan akibat kabut asap pekat.

"Para pandu kapal menunda gerakan kapal-kapal bertonase untuk sementara waktu," kata Kepala Seksi Lalu lintas pelayaran Kantor Kesyahbandaran dan Otoritas Pelabuhan (KSOP) Kelas II Palembang, Andriawan, kepada Antara.

Menurutnya KSOP masih memberlakukan aturan genap - ganjil untuk kapal bertonase masuk dan keluar dari wilayah Sungai Musi Palembang, namun batas waktu hanya pada rentang pukul 06.00 - 10.00 WIB.

Sementara kapal-kapal speedboat pembawa barang dari luar Palembang yang sandar di Dermaga 16 ilir Palembang juga menunda keberangkatannya akibat jarak pandang membahayakan.


"Biasanya pukul 07.00 WIB kami sudah keluar dari Palembang, tapi sampai pukul 08.30 WIB belum bisa keluar karena bahaya sekali, apalagi jam 07.00 - 09.00 WIB itu ramai-ramainya kapal kecil," kata salah seorang nahkoda speedboat, Pardi.


Kabut asap tersebut, kata dia, mengakibatkan omzet pendapatan menurun 20%.

"Penumpang tidak ada yang mau naik karena mereka ada yang takut, ya jadinya kurang penumpang," tambahnya.

Kalimat yang dipilih sebagai ringkasan pada artikel berita 7:

Keterangan :

 : Kalimat terpilih sebagai ringkasan oleh pakar

 : Kalimat terpilih sebagai ringkasan oleh sistem

 : Kalimat terpilih sebagai ringkasan oleh sistem dan pakar

Sejumlah sekolah di Kota Palembang, Sumatera Selatan, memutuskan memulangkan siswa-siswa mereka lantaran kabut asap semakin tebal menyelimuti kota tersebut pada Senin (14/10) pagi.

Kepala SMP Negeri 7 Palembang, Siti Zubaida, mengatakan keputusan pemulangan ditempuh sesuai dengan instruksi Dinas Pendidikan Kota Palembang.

"Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan berdampak buruk terhadap siswa, oleh karenanya atas instruksi Kadiknas Kota Palembang melalui pesan WA Grup meminta siswa dipulangkan dan belajar di rumah masing-masing saja," jelas Siti kepada radio Elshinta.

Hal ini diamini Kepala Dinas Pendidikan Kota Palembang, Ahmad Zulinto, yang menyampaikan surat edaran ke semua sekolah.

"Hari ini seluruh TK hingga SMP negeri dan swasta sederajat diliburkan, untuk besok dan seterusnya akan diberikan edaran lebih lanjut," kata Ahmad Zulinto kepada kantor berita Antara.

Menurutnya, kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan, kegiatan belajar mengajar di tingkat TK sampai SMP akan tetap diliburkan.

Akan tetapi, sebagaimana dipaparkan Kepala Dinas Pendidikan Sumatera Selatan, Widodo, kegiatan belajar mengajar di daerah yang tidak terdampak kabut asap tetap berlangsung.

"Daerah yang tidak terdampak kabut asap tetap normal tetap belajar, untuk daerah yang terkategori sedang tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker, bagi daerah terkategori parah maka siswa diberikan tugas dengan memaksimalkan kelas daring," kata Widodo kepada Antara.

Hal itu belakangan dibenarkan Agus Wibowo, selaku Kepala Pusat Data, Informasi dan Humas Badan Nasional Penanggulangan Bencana (BNPB).

"Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud, TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers.

Sejumlah warga Palembang, Sumatera Selatan, mengeluhkan kabut asap pada Senin (14/10) tergolong parah.

Bahkan jarak pandang hanya 10 meter.

Amelia, warga Kelurahan Bukit Sangkal, Kecamatan Kalidoni, mengatakan dirinya terkejut ketika akan ke luar rumah sekitar pukul 06.30 WIB untuk mengantar anaknya ke sekolah mendapati kabut sedemikian pekat.

"Saya terkejut, kenapa gelap ini.

Kemarin-kemarin ada kabut asap, tapi tidak separah hari ini," kata dia sebagaimana dikutip Antara.

Keluhan atas kondisi ini juga diungkapkan Tina, seorang guru senam di sebuah tempat kebugaran.

"Saya selalu ke luar rumah pukul 06.00 WIB karena ada jadwal senam, sempat terkejut juga karena jarak pandang hanya 10 meter.

Sangat terasa, apalagi saya pakai sepeda motor," kata dia.

Kepala Seksi Observasi dan Informasi Stasiun Meteorologi SMB II Palembang, Bambang Beny Setiaji, mengatakan kabut tersebut bercampur asap kiriman dari wilayah Kabupaten Ogan Komering Ilir (OKI) yang berada sebelah tenggara Kota Palembang.

Kabut asap di Kota Palembang semakin parah dalam sepekan terakhir akibat dampak kebakaran hutan dan lahan di sejumlah kabupaten.

Berdasarkan data Badan Penanggulangan Bencana Daerah Provinsi Sumsel yang bersumber dari Satelit Lapan disebutkan jumlah titik panas pada Senin (14/10) mencapai 732 titik, dengan titik panas terbanyak di Kabupaten Ogan Komering Ilir yang berjumlah 437 titik.

Sebelumnya, pada Jumat (11/10), titik panas berjumlah 417 titik.

Kepala Bidang Kedaruratan Badan Penanggulangan Bencana Daerah Provinsi Sumsel, Ansori, mengatakan titik panas terbanyak terpantau di Kabupaten Ogan Komering Ilir sehingga fokus pemadaman difokuskan di wilayah tersebut.

"Kami selalu lakukan waterbombing (pemadaman dari udara), setiap hari mengerahkan lima unit helikopter.

Kebakaran di OKI ini memang sulit dipadamkan karena terjadi di kawasan gambut, dan akses darat yang terbatas.

belum lagi jika terbakar, asapnya mengarah ke Palembang," kata Ansori.

Sementara itu, aktivitas kapal bertonase di Sungai Musi, Kota Palembang, dihentikan akibat kabut asap pekat.

"Para pandu kapal menunda gerakan kapal-kapal bertonase untuk sementara waktu," kata Kepala Seksi Lalu lintas pelayaran Kantor Kesyahbandaran dan Otoritas Pelabuhan (KSOP) Kelas II Palembang, Andriawan, kepada Antara.

Menurutnya KSOP masih memberlakukan aturan genap - ganjil untuk kapal bertonase masuk dan keluar dari wilayah Sungai Musi Palembang, namun batas waktu hanya pada rentang pukul 06.00 - 10.00 WIB.

Sementara kapal-kapal speedboat pembawa barang dari luar Palembang yang sandar di Dermaga 16 ilir Palembang juga menunda keberangkatannya akibat jarak pandang membahayakan.

"Biasanya pukul 07.00 WIB kami sudah keluar dari Palembang, tapi sampai pukul 08.30 WIB belum bisa keluar karena bahaya sekali, apalagi jam 07.00 - 09.00 WIB itu ramai-ramainya kapal kecil," kata salah seorang nahkoda speedboat, Pardi.

Kabut asap tersebut, kata dia, mengakibatkan omzet pendapatan menurun 20%.

"Penumpang tidak ada yang mau naik karena mereka ada yang takut, ya jadinya kurang penumpang," tambahnya.

B.7 Hasil ringkasan artikel berita untuk evaluasi terburuk dengan *compression rate* 30%

Artikel asli berita 1:

Kerusakan lingkungan disebut Badan Nasional Penanggulangan Bencana (BNPB) sebagai salah satu pemicu banjir dan longsor di Bengkulu.

Bencana alam yang terjadi saat ini dianggap sebagai yang terburuk di Bengkulu dalam belasan tahun terakhir.

Walau terdapat fakta pembukaan pertambangan batu bara dan perkebunan kelapa sawit di beberapa titik di provinsi itu, sejumlah kalangan menyangkal aktivitas tersebut memicu banjir.

Sirajon, warga Desa Arga Mulya, Bengkulu Selatan, mengambil langkah seribu saat arus deras air masuk ke rumahnya yang berjarak 50 meter dari bibir sungai, akhir pekan lalu.

Banjir ini disebut Sirajon yang terbesar yang pernah melanda desanya.

Menurutnya, kejadian ini berkaitan dengan perkebunan sawit yang belakangan di buka sekitar dua kilometer dari Arga Mulya.

"Saya di desa ini sejak 2003.

Sebelumnya pernah terjadi tahun 2012, tapi hanya menggenangi jalan," kata Sirajon saat dihubungi, Selasa (30/04).

"Dampak hujan mungkin karena intensitas hujan tinggi, tapi di sini juga ada perkebunan sawit sehingga hutan untuk menghambat air sudah tidak seperti dulu lagi," tuturnya.

Banjir ini juga menimpa Yessi Ameilia Safitri, warga kelurahan Tanjung Agung, yang berjarak sekitar 230 kilometer dari Arga Mulya.

Ia berkata, selama ini banjir tak pernah masuk ke rumahnya yang lebih tinggi ketimbang jalan raya.

"Kalau rumah saya kena banjir, berarti kelurahan lain sudah pasti banjir total.

Di ujung kompleks rumah saya air setinggi pinggang orang dewasa, tidak mungkin dilewati lagi."

"Saya sekeluarga mengungsi.

Yang bisa dilakukan cuma lari dari rumah dan membawa barang yang bisa ditarik," ucap Yessi.

Menurut data BNPB, banjir melanda sembilan kabupaten di Bengkulu.

Setidaknya 29 orang meninggal dan 13 lainnya hilang akibat banjir yang disusul longsor.

Bencana ini berdampak bagi sekitar 13 ribu warga Bengkulu.

Sebanyak 12 ribu orang di antaranya terpaksa mengungsi.

Juru bicara BNPB, Sutopo Purwo Nugroho, menyebut intensitas tinggi selama beberapa hari di Bengkulu tak setara dengan kapasitas sungai yang ada.

Namun ia menggarisbawahi pula kerusakan lingkungan di sekitar sungai yang memperburuk dampak hujan besar.

"Degradasi hutan, kerusakan lahan, degradasi daerah aliran sungai tinggi.

Dipicu curah hujan maka terjadilah bencana yang masif," kata Sutopo dalam jumpa pers di Jakarta.

"Di Bengkulu Tengah, dampak kerusakan lingkungan dan kerusakan hutan meningkat.

Ada delapan titik yang akan mengubah bentang alam," tuturnya.

Sebelumnya Menteri Energi dan Sumber Daya Mineral Ignasius Jonan menyebut delapan izin tambang batu bara memang dikeluarkan pemerintah setempat beberapa tahun terakhir.

Kantor berita Antara menyebut delapan pemegang izin batu bara di Bengkulu itu adalah PT Bengkulu Bio Energi, PT Kusuma Raya Utama, PT Bara Mega Quantum, PT Inti Bara Perdana.

Empat perusahaan lainnya adalah PT Danau Mas Hitam, PT Ratu Samban Mining, PT Griya Pat Petulai, dan PT Cipta Buana Seraya.

Ada pula satu izin perkebunan kelapa sawit yang dikeluarkan untuk PT Agriandalas yang berada di daerah tangkapan air Sungai Bengkulu.

Di sisi lain, pejabat Kementerian Lingkungan Hidup dan Kehutanan (KLHK) menyangkal aktivitas beberapa perusahaan ini menurunkan kemampuan air tanah menyerap air hujan.

"Posisi izin itu di hulu, tidak terkait langsung dengan sungai.

Jadi banjir ini adalah limpasan, curah hujan melebihi kapasitas sungai," kata Kasubdit Monitoring dan Evaluasi DAS KLHK, Ernawati.

"Izin dikeluarkan jauh dari sungai.

Kajian pada saat aliran air melimpah kami bedakan dengan analisis dampak lingkungannya," ujarnya di Jakarta.

Bagaimanapun, dampak banjir dan longsor di Bengkulu masih terus dirasakan warga setempat, meski air berangsur surut.

BNPB menyebut pasukan evakuasi dan bantuan logistik masih sulit dikerahkan ke dua kecamatan yang terisolasi karena longsor, yaitu Pagar Jati dan Merigi Sakti.

Kawasan Gunung Bungkok yang berada di sekitar dua kecamatan itu tercatat sebagai kawasan longsor dengan jumlah korban meninggal terbanyak.

Badan Meteorologi, Klimatologi, dan Geofisika memperkirakan curah hujan tinggi masih akan melanda Bengkulu dan beberapa daerah lain seperti Aceh, Sumatera Selatan, Sulawesi Utara, Papua, dan Papua Barat.

Masyarakat diimbau untuk tetap waspada, karena banjir dan gerakan tanah berpotensi terjadi lagi pada Mei mendatang.

Kalimat yang dipilih sebagai ringkasan pada artikel berita 1:

Keterangan :

 : Kalimat terpilih sebagai ringkasan oleh pakar

 : Kalimat terpilih sebagai ringkasan oleh sistem

: Kalimat terpilih sebagai ringkasan oleh sistem dan pakar

Kerusakan lingkungan disebut Badan Nasional Penanggulangan Bencana (BNPB) sebagai salah satu pemicu banjir dan longsor di Bengkulu.

Bencana alam yang terjadi saat ini dianggap sebagai yang terburuk di Bengkulu dalam belasan tahun terakhir.

Walau terdapat fakta pembukaan pertambangan batu bara dan perkebunan kelapa sawit di beberapa titik di provinsi itu, sejumlah kalangan menyangkal aktivitas tersebut memicu banjir.

Sirajon, warga Desa Arga Mulya, Bengkulu Selatan, mengambil langkah seribu saat arus deras air masuk ke rumahnya yang berjarak 50 meter dari bibir sungai, akhir pekan lalu.

Banjir ini disebut Sirajon yang terbesar yang pernah melanda desanya.

Menurutnya, kejadian ini berkaitan dengan perkebunan sawit yang belakangan di buka sekitar dua kilometer dari Arga Mulya.

"Saya di desa ini sejak 2003.

Sebelumnya pernah terjadi tahun 2012, tapi hanya menggenangi jalan," kata Sirajon saat dihubungi, Selasa (30/04).

"Dampak hujan mungkin karena intensitas hujan tinggi, tapi di sini juga ada perkebunan sawit sehingga hutan untuk menghambat air sudah tidak seperti dulu lagi," tuturnya.

Banjir ini juga menimpa Yessi Ameilia Safitri, warga kelurahan Tanjung Agung, yang berjarak sekitar 230 kilometer dari Arga Mulya.

Ia berkata, selama ini banjir tak pernah masuk ke rumahnya yang lebih tinggi ketimbang jalan raya.

"Kalau rumah saya kena banjir, berarti kelurahan lain sudah pasti banjir total.

Di ujung kompleks rumah saya air setinggi pinggang orang dewasa, tidak mungkin dilewati lagi."

"Saya sekeluarga mengungsi.

Yang bisa dilakukan cuma lari dari rumah dan membawa barang yang bisa ditarik," ucap Yessi.

Menurut data BNPB, banjir melanda sembilan kabupaten di Bengkulu.

Setidaknya 29 orang meninggal dan 13 lainnya hilang akibat banjir yang disusul longsor.

Bencana ini berdampak bagi sekitar 13 ribu warga Bengkulu.

Sebanyak 12 ribu orang di antaranya terpaksa mengungsi.

Juru bicara BNPB, Sutopo Purwo Nugroho, menyebut intensitas tinggi selama beberapa hari di Bengkulu tak setara dengan kapasitas sungai yang ada.

Namun ia menggarisbawahi pula kerusakan lingkungan di sekitar sungai yang memperburuk dampak hujan besar.

"Degradasi hutan, kerusakan lahan, degradasi daerah aliran sungai tinggi.

Dipicu curah hujan maka terjadilah bencana yang masif," kata Sutopo dalam jumpa pers di Jakarta.

"Di Bengkulu Tengah, dampak kerusakan lingkungan dan kerusakan hutan meningkat.

Ada delapan titik yang akan mengubah bentang alam," tuturnya.

Sebelumnya Menteri Energi dan Sumber Daya Mineral Ignasius Jonan menyebut delapan izin tambang batu bara memang dikeluarkan pemerintah setempat beberapa tahun terakhir.

Kantor berita Antara menyebut delapan pemegang izin batu bara di Bengkulu itu adalah PT Bengkulu Bio Energi, PT Kusuma Raya Utama, PT Bara Mega Quantum, PT Inti Bara Perdana.

Empat perusahaan lainnya adalah PT Danau Mas Hitam, PT Ratu Samban Mining, PT Griya Pat Petulai, dan PT Cipta Buana Seraya.

Ada pula satu izin perkebunan kelapa sawit yang dikeluarkan untuk PT Agriandalas yang berada di daerah tangkapan air Sungai Bengkulu.

Di sisi lain, pejabat Kementerian Lingkungan Hidup dan Kehutanan (KLHK) menyangkal aktivitas beberapa perusahaan ini menurunkan kemampuan air tanah menyerap air hujan.

"Posisi izin itu di hulu, tidak terkait langsung dengan sungai.

Jadi banjir ini adalah limpasan, curah hujan melebihi kapasitas sungai," kata Kasubdit Monitoring dan Evaluasi DAS KLHK, Ernawati.

"Izin dikeluarkan jauh dari sungai.

Kajian pada saat aliran air melimpah kami bedakan dengan analisis dampak lingkungannya," ujarnya di Jakarta.

Bagaimanapun, dampak banjir dan longsor di Bengkulu masih terus dirasakan warga setempat, meski air berangsur surut.

BNPB menyebut pasukan evakuasi dan bantuan logistik masih sulit dikerahkan ke dua kecamatan yang terisolasi karena longsor, yaitu Pagar Jati dan Merigi Sakti.

Kawasan Gunung Bungkok yang berada di sekitar dua kecamatan itu tercatat sebagai kawasan longsor dengan jumlah korban meninggal terbanyak.

Badan Meteorologi, Klimatologi, dan Geofisika memperkirakan curah hujan tinggi masih akan melanda Bengkulu dan beberapa daerah lain seperti Aceh, Sumatera Selatan, Sulawesi Utara, Papua, dan Papua Barat.

Masyarakat diimbau untuk tetap waspada, karena banjir dan gerakan tanah berpotensi terjadi lagi pada Mei mendatang.

B.8 Hasil ringkasan artikel berita untuk evaluasi terbaik dengan *compression rate 30%*

Artikel asli berita 7:

Sejumlah sekolah di Kota Palembang, Sumatera Selatan, memutuskan memulangkan siswa-siswa mereka lantaran kabut asap semakin tebal menyelimuti kota tersebut pada Senin (14/10) pagi.

Kepala SMP Negeri 7 Palembang, Siti Zubaida, mengatakan keputusan pemulangan ditempuh sesuai dengan instruksi Dinas Pendidikan Kota Palembang.

"Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan berdampak buruk terhadap siswa, oleh karenanya atas instruksi Kadiknas Kota Palembang melalui pesan WA Grup meminta siswa dipulangkan dan belajar di rumah masing-masing saja," jelas Siti kepada radio Elshinta.

Hal ini diamini Kepala Dinas Pendidikan Kota Palembang, Ahmad Zulinto, yang menyampaikan surat edaran ke semua sekolah.

"Hari ini seluruh TK hingga SMP negeri dan swasta sederajat diliburkan, untuk besok dan seterusnya akan diberikan edaran lebih lanjut," kata Ahmad Zulinto kepada kantor berita Antara.

Menurutnya, kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan, kegiatan belajar mengajar di tingkat TK sampai SMP akan tetap diliburkan.

Akan tetapi, sebagaimana dipaparkan Kepala Dinas Pendidikan Sumatera Selatan, Widodo, kegiatan belajar mengajar di daerah yang tidak terdampak kabut asap tetap berlangsung.

"Daerah yang tidak terdampak kabut asap tetap normal tetap belajar, untuk daerah yang terkategori sedang tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker, bagi daerah terkategori parah maka siswa diberikan tugas dengan memaksimalkan kelas daring," kata Widodo kepada Antara.

Hal itu belakangan dibenarkan Agus Wibowo, selaku Kepala Pusat Data, Informasi dan Humas Badan Nasional Penanggulangan Bencana (BNPB).

"Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud, TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers.

Sejumlah warga Palembang, Sumatera Selatan, mengeluhkan kabut asap pada Senin (14/10) tergolong parah.

Bahkan jarak pandang hanya 10 meter.

Amelia, warga Kelurahan Bukit Sangkal, Kecamatan Kalidoni, mengatakan dirinya terkejut ketika akan ke luar rumah sekitar pukul 06.30 WIB untuk mengantar anaknya ke sekolah mendapati kabut sedemikian pekat.

"Saya terkejut, kenapa gelap ini.

Kemarin-kemarin ada kabut asap, tapi tidak separah hari ini," kata dia sebagaimana dikutip Antara.

Keluhan atas kondisi ini juga diungkapkan Tina, seorang guru senam di sebuah tempat kebugaran.

"Saya selalu ke luar rumah pukul 06.00 WIB karena ada jadwal senam, sempat terkejut juga karena jarak pandang hanya 10 meter.

Sangat terasa, apalagi saya pakai sepeda motor," kata dia.

Kepala Seksi Observasi dan Informasi Stasiun Meteorologi SMB II Palembang, Bambang Beny Setiaji, mengatakan kabut tersebut bercampur asap kiriman dari wilayah Kabupaten Ogan Komering Ilir (OKI) yang berada sebelah tenggara Kota Palembang.

Kabut asap di Kota Palembang semakin parah dalam sepekan terakhir akibat dampak kebakaran hutan dan lahan di sejumlah kabupaten.

Berdasarkan data Badan Penanggulangan Bencana Daerah Provinsi Sumsel yang bersumber dari Satelit Lapan disebutkan jumlah titik panas pada Senin (14/10) mencapai 732 titik, dengan titik panas terbanyak di Kabupaten Ogan Komering Ilir yang berjumlah 437 titik.

Sebelumnya, pada Jumat (11/10), titik panas berjumlah 417 titik.

Kepala Bidang Kedaruratan Badan Penanggulangan Bencana Daerah Provinsi Sumsel, Ansori, mengatakan titik panas terbanyak terpantau di Kabupaten Ogan Komering Ilir sehingga fokus pemadaman difokuskan di wilayah tersebut.

"Kami selalu lakukan waterbombing (pemadaman dari udara), setiap hari mengerahkan lima unit helikopter.

Kebakaran di OKI ini memang sulit dipadamkan karena terjadi di kawasan gambut, dan akses darat yang terbatas.

belum lagi jika terbakar, asapnya mengarah ke Palembang," kata Ansori.

Sementara itu, aktivitas kapal bertonase di Sungai Musi, Kota Palembang, dihentikan akibat kabut asap pekat.

"Para pandu kapal menunda gerakan kapal-kapal bertonase untuk sementara waktu," kata Kepala Seksi Lalu lintas pelayaran Kantor Kesyahbandaran dan Otoritas Pelabuhan (KSOP) Kelas II Palembang, Andriawan, kepada Antara.

Menurutnya KSOP masih memberlakukan aturan genap - ganjil untuk kapal bertonase masuk dan keluar dari wilayah Sungai Musi Palembang, namun batas waktu hanya pada rentang pukul 06.00 - 10.00 WIB.

Sementara kapal-kapal speedboat pembawa barang dari luar Palembang yang sandar di Dermaga 16 ilir Palembang juga menunda keberangkatannya akibat jarak pandang membahayakan.


"Biasanya pukul 07.00 WIB kami sudah keluar dari Palembang, tapi sampai pukul 08.30 WIB belum bisa keluar karena bahaya sekali, apalagi jam 07.00 - 09.00 WIB itu ramai-ramainya kapal kecil," kata salah seorang nahkoda speedboat, Pardi.

Kabut asap tersebut, kata dia, mengakibatkan omzet pendapatan menurun 20%.

"Penumpang tidak ada yang mau naik karena mereka ada yang takut, ya jadinya kurang penumpang," tambahnya.

Kalimat yang dipilih sebagai ringkasan pada artikel berita 7:

Keterangan :

 : Kalimat terpilih sebagai ringkasan oleh pakar

 : Kalimat terpilih sebagai ringkasan oleh sistem

 : Kalimat terpilih sebagai ringkasan oleh sistem dan pakar

Sejumlah sekolah di Kota Palembang, Sumatera Selatan, memutuskan memulangkan siswa-siswa mereka lantaran kabut asap semakin tebal menyelimuti kota tersebut pada Senin (14/10) pagi.

Kepala SMP Negeri 7 Palembang, Siti Zubaida, mengatakan keputusan pemulangan ditempuh sesuai dengan instruksi Dinas Pendidikan Kota Palembang.

"Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan berdampak buruk terhadap siswa, oleh karenanya atas instruksi Kadiknas Kota Palembang melalui pesan WA Grup meminta siswa dipulangkan dan belajar di rumah masing-masing saja," jelas Siti kepada radio Elshinta.

Hal ini diamini Kepala Dinas Pendidikan Kota Palembang, Ahmad Zulinto, yang menyampaikan surat edaran ke semua sekolah.

"Hari ini seluruh TK hingga SMP negeri dan swasta sederajat diliburkan, untuk besok dan seterusnya akan diberikan edaran lebih lanjut," kata Ahmad Zulinto kepada kantor berita Antara.

Menurutnya, kalau kualitas udara Palembang masih buruk dalam beberapa hari ke depan, kegiatan belajar mengajar di tingkat TK sampai SMP akan tetap diliburkan.

Akan tetapi, sebagaimana dipaparkan Kepala Dinas Pendidikan Sumatera Selatan, Widodo, kegiatan belajar mengajar di daerah yang tidak terdampak kabut asap tetap berlangsung.

"Daerah yang tidak terdampak kabut asap tetap normal tetap belajar, untuk daerah yang terkategori sedang tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker, bagi daerah terkategori parah maka siswa diberikan tugas dengan memaksimalkan kelas daring," kata Widodo kepada Antara.

Hal itu belakangan dibenarkan Agus Wibowo, selaku Kepala Pusat Data, Informasi dan Humas Badan Nasional Penanggulangan Bencana (BNPB).

"Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat paud, TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers.

Sejumlah warga Palembang, Sumatera Selatan, mengeluhkan kabut asap pada Senin (14/10) tergolong parah.

Bahkan jarak pandang hanya 10 meter.

Amelia, warga Kelurahan Bukit Sangkal, Kecamatan Kalidoni, mengatakan dirinya terkejut ketika akan ke luar rumah sekitar pukul 06.30 WIB untuk mengantar anaknya ke sekolah mendapati kabut sedemikian pekat.

"Saya terkejut, kenapa gelap ini.

Kemarin-kemarin ada kabut asap, tapi tidak separah hari ini," kata dia sebagaimana dikutip Antara.

Keluhan atas kondisi ini juga diungkapkan Tina, seorang guru senam di sebuah tempat kebugaran.

"Saya selalu ke luar rumah pukul 06.00 WIB karena ada jadwal senam, sempat terkejut juga karena jarak pandang hanya 10 meter.

Sangat terasa, apalagi saya pakai sepeda motor," kata dia.

Kepala Seksi Observasi dan Informasi Stasiun Meteorologi SMB II Palembang, Bambang Beny Setiaji, mengatakan kabut tersebut bercampur asap kiriman dari wilayah Kabupaten Ogan Komering Ilir (OKI) yang berada sebelah tenggara Kota Palembang.

Kabut asap di Kota Palembang semakin parah dalam sepekan terakhir akibat dampak kebakaran hutan dan lahan di sejumlah kabupaten.

Berdasarkan data Badan Penanggulangan Bencana Daerah Provinsi Sumsel yang bersumber dari Satelit Lapan disebutkan jumlah titik panas pada Senin (14/10) mencapai 732 titik, dengan titik panas terbanyak di Kabupaten Ogan Komering Ilir yang berjumlah 437 titik.

Sebelumnya, pada Jumat (11/10), titik panas berjumlah 417 titik.

Kepala Bidang Kedaruratan Badan Penanggulangan Bencana Daerah Provinsi Sumsel, Ansori, mengatakan titik panas terbanyak terpantau di Kabupaten Ogan Komering Ilir sehingga fokus pemadaman difokuskan di wilayah tersebut.

"Kami selalu lakukan waterbombing (pemadaman dari udara), setiap hari mengerahkan lima unit helikopter.

Kebakaran di OKI ini memang sulit dipadamkan karena terjadi di kawasan gambut, dan akses darat yang terbatas.

belum lagi jika terbakar, asapnya mengarah ke Palembang," kata Ansori.

Sementara itu, aktivitas kapal bertonase di Sungai Musi, Kota Palembang, dihentikan akibat kabut asap pekat.

"Para pandu kapal menunda gerakan kapal-kapal bertonase untuk sementara waktu," kata Kepala Seksi Lalu lintas pelayaran Kantor Kesyahbandaran dan Otoritas Pelabuhan (KSOP) Kelas II Palembang, Andriawan, kepada Antara.

Menurutnya KSOP masih memberlakukan aturan genap - ganjil untuk kapal bertonase masuk dan keluar dari wilayah Sungai Musi Palembang, namun batas waktu hanya pada rentang pukul 06.00 - 10.00 WIB.

Sementara kapal-kapal speedboat pembawa barang dari luar Palembang yang sandar di Dermaga 16 ilir Palembang juga menunda keberangkatannya akibat jarak pandang membahayakan.

"Biasanya pukul 07.00 WIB kami sudah keluar dari Palembang, tapi sampai pukul 08.30 WIB belum bisa keluar karena bahaya sekali, apalagi jam 07.00 - 09.00 WIB itu ramai-ramainya kapal kecil," kata salah seorang nahkoda speedboat, Pardi.

Kabut asap tersebut, kata dia, mengakibatkan omzet pendapatan menurun 20%.

"Penumpang tidak ada yang mau naik karena mereka ada yang takut, ya jadinya kurang penumpang," tambahnya.