

PERINGKASAN ARTIKEL BERBAHASA INDONESIA MENGGUNAKAN TEXTRANK DENGAN PEMBOBOTAN BM25

Yurdha Fadhila Hernawan ^{*1}, Putra Pandu Adikara ², Randy Cahya Wihandika ³

^{1,2,3}Universitas Brawijaya, Malang

Email: ¹yurdhafadhila@student.ub.ac.id, ²adikara.putra@ub.ac.id, ³rendicahya@ub.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 20 Juli 2020, diterima untuk diterbitkan: 27 Januari 2022)

Abstrak

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, penelitian ini berhasil dilakukan dengan kualitas ringkasan terbaik didapatkan pada saat penggunaan *compression rate* sebesar 30% dengan nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,552; 0,552; dan 0,552.

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

AUTOMATIC SUMMARIZATION OF INDONESIAN ARTICLES USING TEXTRANK WITH BM25 WEIGHTING

Abstract

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, every information accessed is not 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, this research is successfully carried out with 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*

1. PENDAHULUAN

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. 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).

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).

Peringkasan ekstraktif telah dilakukan oleh (Barrios et al., 2016) dengan menggunakan metode *TextRank*. *TextRank* merupakan metode pemeringkatan kalimat berbasis graf. Untuk setiap kalimat pada dokumen direpresentasikan sebagai *node* dan hubungan antara satu kalimat dengan kalimat lainnya direpresentasikan sebagai *edges*. Metode yang digunakan untuk mendapatkan hubungan antara satu kalimat dengan kalimat lainnya tersebut adalah dengan mencari nilai kemiripan (*similarity*). Nilai kemiripan ini bisa didapatkan dengan melihat kata yang sama yang terdapat antara dua kalimat. Jenis nilai *similarity* yang digunakan pada penelitian (Barrios et al., 2016) adalah *Longest Common Substring*, *cosine distance*, BM25, dan BM25+ dengan bobot pemeringkatan *PageRank*.

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

Untuk setiap hasil *clustering* yang telah didapatkan, peringkasan dilakukan dengan metode *TextRank* dengan menggunakan nilai kemiripan (*similarity*) BM25. Setelah merepresentasikan dokumen dalam bentuk *TextRank*, kalimat diperingkatkan berdasarkan bobot *PageRank*. Kalimat yang terpilih berdasarkan peringkat 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)*.

Masalah utama yang muncul setelah melakukan peringkasan adalah kualitas hasil peringkasan. Ringkasan yang dibuat harus dapat memenuhi dan merepresentasikan semua konteks dan informasi penting pada sebuah dokumen. Salah satu parameter pada proses pembuatan ringkasan adalah panjang ringkasan yang dibuat. Untuk melakukan peringkasan ekstraktif dengan memilih sejumlah kalimat penting pada dokumen, jumlah kalimat yang dipilih sebagai ringkasan sangat mempengaruhi kualitas ringkasan yang dihasilkan. Jika jumlah kalimat yang dipilih terlalu sedikit maka banyak informasi penting yang dibuang, sedangkan jika kalimat yang dipilih terlalu banyak maka banyak informasi yang tidak penting yang dijadikan sebagai ringkasan.

Pada proses pembentukan ringkasan, panjang ringkasan tersebut didapatkan dari bobot *compression rate* (rasio kompresi). Rasio kompresi merupakan besaran jumlah kalimat yang dijadikan ringkasan berdasarkan jumlah kalimat yang terdapat pada suatu dokumen. Contoh kasus untuk dokumen yang berisikan 10 kalimat dengan bobot *compression rate* sebesar 20%, maka panjang ringkasan yang dihasilkan adalah sebanyak 2 kalimat. Nilai *compression rate* yang digunakan untuk mendapatkan ringkasan tidak lebih besar dari 50% (kurang dari setengah panjang dokumen), oleh sebab itu pengujian terhadap hasil ringkasan berdasarkan *compression rate* perlu dilakukan untuk mengetahui bahwa hasil ringkasan dapat diterima.

Penelitian ini dilakukan untuk mendapatkan ringkasan secara otomatis dengan objek artikel berita *online* berbahasa Indonesia menggunakan metode pemeringkatan *TextRank* dan BM25 sebagai fungsi *similarity*. Metode *TextRank* cocok digunakan untuk dokumen terstruktur seperti artikel berita yang mana

setiap kalimat memiliki beberapa informasi yang relevan dan memiliki sedikit redundansi informasi (Barrios et al., 2016). 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. Pengujian hasil ringkasan dilakukan dengan mencoba beberapa nilai *compression rate*.

2. METODE PENELITIAN

Peringkasan teks dapat diartikan sebagai proses pengambilan dokumen sebagai masukan dan dokumen lebih pendek yang hanya berisikan konten penting sebagai keluaran. Selain dikategorikan berdasarkan jumlah dokumen dan tipe peringkasan, pemeringkatan teks juga dapat dikategorikan berdasarkan topik ringkasan yang dihasilkan. Ringkasan umum (*generic*) akan menghasilkan ringkasan tanpa mempertimbangkan topik yang lebih spesifik, sedangkan pemeringkatan berdasarkan *query-relevant* hanya akan menghasilkan ringkasan yang sesuai dengan *query* atau kata kunci yang diberikan (Jackson and Moulinier, 2004).

Peringkasan teks juga dapat dikatakan sebagai solusi dari masalah *information overload* dan proses reduksi data secara cepat. Beberapa keunggulan yang didapatkan dari peringkasan otomatis ialah: hasil ringkasan yang dihasilkan tidak lebih bias dibandingkan dengan ringkasan yang dihasilkan manusia, peringkasan teks sangat diperlukan untuk sistem otomatis tanya-jawab, efisiensi durasi yang digunakan untuk membaca ringkasan dibandingkan dengan membaca seluruh dokumen, juga memudahkan para peneliti dalam proses pemilihan artikel saat melakukan penelitian (Mohd, Jan and Shah, 2020).

Salah satu penelitian yang menjadi acuan 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 Kazakhstan, 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 dalam sebuah dokumen 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 *key-words*, dengan nilai rata-rata untuk nilai *similarity* menggunakan *content overlap*, *LongestCommonSubstring*, dan BM25 secara berurutan adalah 0,180; 0,175; dan 0,169.

Peringkasan dokumen berbahasa Indonesia telah dilakukan oleh (Najibullah, 2015) dengan menggunakan metode klasifikasi Naïve bayes. Metode naïve bayes digunakan untuk menentukan apakah kalimat pada dokumen dapat dijadikan sebagai ringkasan atau tidak berdasarkan nilai probabilitas. Dokumen yang diringkas dipecah menjadi potongan-potongan kalimat dan digunakan untuk proses ekstraksi fitur.

Ada 10 fitur yang digunakan pada penelitian tersebut, beberapa diantaranya adalah posisi kalimat, nilai relatif panjang kalimat, nilai rata-rata *tf*, hasil perkalian nilai *tf* dan *idf*, *keyphrase extraction*, dan nilai *similarity* antara kalimat dengan judul dokumen. Penelitian ini menggunakan 100 dokumen berita berbahasa Indonesia yang diambil dari portal berita detik news. 100 dokumen tersebut dibagi menjadi 80% sebagai data latih dan 20% sebagai data uji. Penelitian tersebut menguji nilai *compression rate* dengan menghitung nilai *precision*, *recall*, dan *f-measure*. Penelitian tersebut menghasilkan nilai evaluasi ringkasan yang bagus ketika menggunakan *compression rate* sebesar 30% yaitu dengan nilai *precision*, *recall*, dan *f-measure* secara berurut adalah 0,7385; 0,7195, dan 0,7152.

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 menghitung *Longest Common Substring*, *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*.

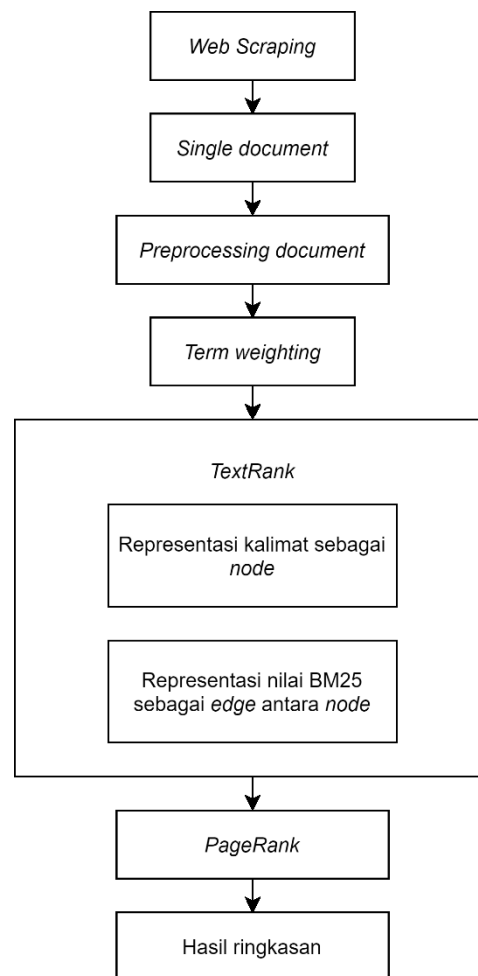
Dokumen yang digunakan pada penelitian (Barrios et al., 2016) adalah *Document Understanding Conference (DUC)* yang berjumlah sebanyak 567 dokumen dengan peringkasan sebanyak 20% dari tiap panjang dokumen. Hasil ringkasan dari percobaan tersebut dievaluasi menggunakan metode *ROUGE-N* dengan nilai terbaik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

Berdasarkan pemaparan tersebut, tahapan penelitian ini dimulai dengan pengumpulan data artikel berita berbahasa Indonesia yang digunakan untuk menguji sistem peringkasan yang dibuat. Selanjutnya perancangan dan implementasi sistem peringkasan otomatis menggunakan metode *TextRank* dengan pembobotan BM25, lalu menguji sistem tersebut dengan membandingkan hasil ringkasan yang didapatkan oleh sistem otomatis dengan hasil ringkasan yang didapatkan oleh *expert* (pakar) yang telah ditentukan.

Artikel berita yang digunakan pada penelitian ini didapatkan menggunakan metode *web scraping*, sehingga data masukan yang dibutuhkan merupakan URL berita portal *online* berbahasa Indonesia. Pada metode *TextRank*, artikel berita direpresentasikan sebagai graf, yang mana setiap kalimat dianggap sebagai *node* dan nilai pembobotan BM25 sebagai *edge* antar *node*. Selanjutnya setiap kalimat (*node*) diperingkatkan berdasarkan nilai *PageRank*, kalimat dengan nilai *PageRank* yang tinggi dianggap sebagai kalimat penting dan dapat merepresentasikan dokumen sebagai ringkasan.

Pada Gambar 1 hasil *web scraping* setelah memasukkan URL berita adalah sebuah teks panjang yang merupakan sebuah dokumen berita (*single document*) yang akan diringkas. Dokumen tersebut dibersihkan dengan menghilangkan kata *stopword*, angka, tanda baca dan mengubah seluruh kalimat menjadi huruf kecil (*lower case*). Selain itu pada proses pembersihan ini dokumen dipisahkan menjadi potongan-potongan kalimat. Proses pembersihan ini disebut juga dengan *preprocessing document*.

Proses *term weighting* digunakan untuk menghitung jumlah kemunculan sebuah kata dalam kalimat (*term frequency*) dan nilai inverse dari jumlah kalimat yang memiliki kemunculan sebuah kata (*inverse document frequency*). Kedua nilai *tf* dan *idf* ini digunakan pada proses perhitungan *similarity* BM25. Selanjutnya dokumen direpresentasikan sebagai graf (*TextRank*) dan menghitung pemeringkatan setiap kalimat (*PageRank*). Kalimat pada dokumen diurutkan berdasarkan nilai *PageRank* terbesar dan dipilih sesuai bobot *compression rate* untuk dijadikan ringkasan.



Gambar 1. Gambaran Umum Proses Peringkasan Otomatis

2.1 Pengumpulan Data

Artikel berita berbahasa Indonesia yang digunakan pada penelitian ini diambil melalui situs berita BBC Indonesia sebanyak 10 berita mengenai bencana alam dan lingkungan. Pertimbangan jumlah berita yang sedikit diambil karena penelitian ini berfokuskan pada penelitian yang mendalam berdasarkan jenis *compression rate* yang digunakan untuk mendapatkan ringkasan.

2.2 TextRank

TextRank merupakan sebuah algoritme berbasis graf yang digunakan untuk menentukan kalimat penting pada suatu dokumen (Mutlu, Sezer and Akcayol, 2020). 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.

2.3 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. Fungsi *similarity* BM25 didapatkan dari perhitungan bobot *tf* dan *idf* pada setiap kata (*term*). Selain itu juga dipengaruhi parameter 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 (1).

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

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

2.4 PageRank

PageRank adalah metode yang digunakan dalam pemeringkatan graf (Mussina, Aubakirov and Trigo, 2018). *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 memiliki nilai *PageRank* yang besar.

Inisialisasi awal nilai *PageRank* tiap kalimat ditentukan secara random mulai dari 0 hingga 1. Lalu sejumlah iterasi dilakukan untuk melakukan update bobot *PageRank* di tiap kalimat. Persamaan *PageRank* dijabarkan pada Persamaan (2).

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

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*.

2.5 Compression Rate

Compression rate digunakan sebagai acuan panjang ringkasan yang dihasilkan sistem peringkasan otomatis. Ringkasan yang bagus seharusnya memiliki nilai *compression rate* yang rendah dan nilai *retention rate* yang tinggi (Ermakova, Cossu and Mothe, 2019). Besar *compression rate* yang digunakan pada penelitian ini adalah 5%, 10%, 20%, dan 30%.

2.6 Evaluasi Hasil Ringkasan

Evaluasi hasil ringkasan yang didapatkan sistem otomatis dilakukan dengan membandingkan hasil ringkasan sistem dengan hasil ringkasan yang didapatkan oleh pakar. Pakar yang berpartisipasi dalam penelitian ini adalah dosen Bahasa Indonesia dari Universitas Brawijaya. Pakar diminta untuk memilih kalimat yang dianggap penting dan dapat dijadikan ringkasan. Perhitungan *precision*, *recall*, dan *F-Measure* digunakan untuk mengetahui kualitas ringkasan otomatis tersebut.

3. IMPLEMENTASI SISTEM

Sistem peringkasan otomatis dibuat dengan menggunakan bahasa pemrograman *python* dan *software* Jupyter Notebook. Untuk mempermudah proses *scraping* dalam mendapatkan dokumen berita, implementasi sistem dibantu dengan menggunakan library BeautifulSoup. Setelah memasukan URL berita BBC Indonesia dan memilih panjang ringkasan berdasarkan jenis *compression rate*, sistem mengembalikan hasil ringkasan berita beserta judul berita yang diringkaskan seperti yang ditunjukkan pada Gambar 2.

```
URL berita BBC Indonesia: https://www.bbc.com/indonesia/indonesia-50038237
Jenis compression rate:
1. 5%
2. 10%
3. 20%
4. 30%
Pilih jenis compression rate: 4

Judul berita:
Asap Palembang: Kabut selimuti ibu kota Sumatera Selatan, siswa sekolah diliburkan

Hasil ringkasan:
```

Sejumlah sekolah di Kota Palembang, Sumatera Selatan, memutuskan memulangkan siswa-siswa mereka lantaran kabut asap semakin tebal menyelimuti kota tersebut. "Pagi ini kami memulangkan siswa karena melihat kabut asap yang tebal dan benda hitam di udara," jelas Siti kepada radio Elshinta. 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 terdampak kabut asap tetap belajar namun jam masuk sekolah diundur dan kami himbau memakai masker, bagi daerah terdampak parah maka siswa diberikan tugas dengan memanfaatkan kelas daring," kata Widodo kepada Antara. "Melalui pesan digital, Kepala Dinas Pendidikan Kota Palembang menginstruksikan kegiatan belajar mengajar di tingkat TK, SD dan SMP negeri dan swasta diliburkan hingga batas yang belum ditentukan," sebut Agus dalam siaran pers. 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 Sumatera Selatan 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. Kepala Bidang Kedaruratan Badan Penanggulangan Bencana Daerah Provinsi Sumsel, Ansori, mengatakan titik panas terbanyak terdapat di Kabupaten Ogan Komering Ilir sehingga fokus pemadaman difokuskan di wilayah tersebut. Sementara itu, aktivitas kapal bertonase di Sungai Musi, Kota Palembang, dihentikan akibat kabut asap pekat.

Gambar 2. Hasil Implementasi

4. HASIL DAN PEMBAHASAN

Tabel 1 menunjukkan nilai evaluasi sepuluh dokumen dengan *compression rate* sebesar 5%. Beberapa dokumen memiliki nilai evaluasi sebesar nol yaitu pada dokumen 4, 5, 6, 7, 8, dan 9, yang mana menjelaskan bahwa tidak ada kalimat ringkasan hasil sistem yang dianggap benar atau sama dengan ringkasan yang diperoleh pakar. Sedangkan untuk dokumen 1, 2, 3, dan 8 memiliki setengah ringkasan yang sama antara hasil sistem dan hasil pakar. Untuk sepuluh dokumen yang diuji nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,233; 0,233; dan 0,233. Nilai evaluasi ini cukup rendah.

Tabel 2 menunjukkan nilai evaluasi dari sepuluh dokumen yang diuji menggunakan nilai *compression rate* sebesar 10%. Berbeda dengan Tabel 1, pengujian dengan menggunakan *compression rate* 10% memiliki lebih sedikit dokumen dengan nilai evaluasi 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 sepuluh dokumen yang diuji nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,267; 0,267; dan 0,267.

Tabel 1. Hasil Pengujian *Compression Rate* 5%

Dokumen	Precision	Recall	F-Measure
1	0,5	0,5	0,5
2	0,5	0,5	0,5
3	0,5	0,5	0,5
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0,5	0,5	0,5
9	0	0	0
10	0,333	0,333	0,333
Rata - rata	0,233	0,233	0,233

Tabel 3 menunjukkan nilai dari evaluasi dari sepuluh dokumen yang diuji menggunakan nilai *compression rate* sebesar 20%. Pengujian ini tidak lagi mendapatkan hasil nilai evaluasi sebesar nol. Nilai evaluasi terendah didapatkan pada dokumen 3,

dokumen 4, dan dokumen 6. Sedangkan untuk evaluasi tertinggi didapatkan pada dokumen 7 yaitu sebesar 0,571; 0,571; dan 0,571. Untuk sepuluh dokumen yang diuji nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,369; 0,369; dan 0,369.

Tabel 2. Hasil Pengujian *Compression Rate* 10%

Dokumen	Precision	Recall	F-Measure
1	0,25	0,25	0,25
2	0,333	0,333	0,333
3	0,333	0,333	0,333
4	0	0	0
5	0,5	0,5	0,5
6	0	0	0
7	0,25	0,25	0,25
8	0,5	0,5	0,5
9	0	0	0
10	0,5	0,5	0,5
Rata - rata	0,267	0,267	0,267

Tabel 3. Hasil Pengujian *Compression Rate* 20%

Dokumen	Precision	Recall	F-Measure
1	0,25	0,25	0,25
2	0,333	0,333	0,333
3	0,333	0,333	0,333
4	0	0	0
5	0,5	0,5	0,5
6	0	0	0
7	0,25	0,25	0,25
8	0,5	0,5	0,5
9	0	0	0
10	0,5	0,5	0,5
Rata - rata	0,267	0,267	0,267

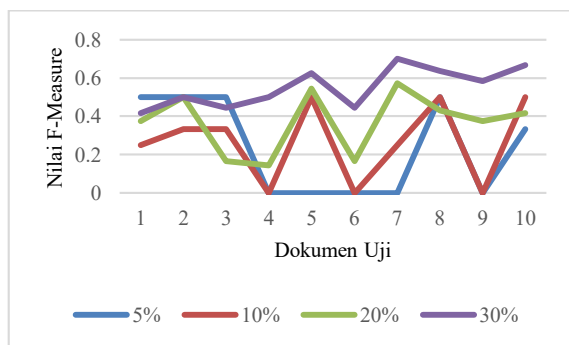
Tabel 4 menunjukkan nilai evaluasi sepuluh dokumen yang diuji menggunakan nilai *compression rate* sebesar 30%. Hasil *precision*, *recall*, dan *f-measure* terendah didapatkan pada dokumen 1, yaitu sebesar 0,417; 0,417; dan 0,417. Sedangkan untuk hasil *precision*, *recall*, dan *f-measure* tertinggi didapatkan pada dokumen 7 yaitu sebesar 0,7; 0,7; dan 0,7. Besar nilai evaluasi untuk penggunaan *compression rate* sebesar 30% tidak lagi mendapatkan nilai nol dan lebih baik dibandingkan dengan penggunaan *compression rate* sebesar 20% menunjukkan bahwa tingkat keberhasilan sistem cukup tinggi pada pengujian ini. Untuk 10 dokumen yang diuji nilai rata-rata *precision*, *recall*, dan *f-measure* secara berturut-turut adalah 0,552; 0,552; dan 0,552.

Tabel 4. Hasil Pengujian *Compression Rate* 30%

Dokumen	Precision	Recall	F-Measure
1	0,417	0,417	0,417
2	0,5	0,5	0,5
3	0,444	0,444	0,444
4	0,5	0,5	0,5
5	0,625	0,625	0,625
6	0,444	0,444	0,444
7	0,7	0,7	0,7
8	0,636	0,636	0,636
9	0,583	0,583	0,583
10	0,667	0,667	0,667
Rata - rata	0,552	0,552	0,552

Perbandingan nilai *f-measure* keempat jenis *compression rate* ditunjukkan pada Gambar 3. Hasil

evaluasi terbaik didapatkan ketika pengujian dilakukan dengan menggunakan bobot *compression rate* 30%. Jika dibandingkan dengan hasil evaluasi penelitian (Najibullah, 2015) yang juga menghasilkan nilai evaluasi terbaik pada penggunaan *compression rate* sebesar 30%, hasil penelitian ini memberikan nilai evaluasi yang cukup baik dengan rata-rata nilai *f-measure* sebesar 0,552. Hasil evaluasi ini tentunya tidak luput dari bias yang didapat dari hasil ringkasan pakar (*expert*).



Gambar 3. Perbandingan Nilai F-Measure 4 Jenis *Compression Rate*

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan implementasi dan evaluasi sistem peringkasan otomatis yang telah berhasil dilakukan dapat disimpulkan bahwa 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.

Setelah membandingkan hasil ringkasan yang didapatkan sistem otomatis dengan hasil ringkasan yang didapatkan dari pakar sebanyak sepuluh 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 0,552; 0,552; dan 0,552.

5.2 Saran

Hal yang dapat diperbaiki dan menjadi masukan untuk penelitian selanjutnya ialah dengan mengikutsertakan judul dokumen pada proses pembentukan ringkasan. Perbedaan dataset (struktur kalimat pada artikel ilmiah berbeda dengan struktur kalimat pada artikel berita) yang digunakan juga dapat mempengaruhi hasil kualitas ringkasan ketika menggunakan fitur judul dokumen saat melakukan ringkasan.

DAFTAR PUSTAKA

ABBASI-GHALEHTAKI, R., KHOTANLOU, H.

AND ESMAEILPOUR, 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>>.

BARRIOS, F., LÓPEZ, F., ARGERICH, L. AND WACHENCHAUZER, R., 2016. Variations of the Similarity Function of TextRank for Automated Summarization. [online] Available at: <<http://arxiv.org/abs/1602.03606>>.

ERMAKOVA, L., COSSU, J.V. AND MOTHE, J., 2019. A survey on evaluation of summarization methods. *Information Processing and Management*, [online] 56(5), pp.1794–1814. Available at: <<https://doi.org/10.1016/j.ipm.2019.04.001>>.

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>>.

JACKSON, P. AND MOULINIER, I., 2004. Natural language processing for online applications: Text retrieval, extraction and categorization. *Terminology Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 10(1), pp.177–179.

MANNING, C.C.; RAGGHAVAN, P.; SCHÜTZE, H., 2009. *An Introduction to Information Retrival*. Cambridge, England: Cambridge University Press.

MOHD, M., JAN, R. AND SHAH, M., 2020. Text document summarization using word embedding. *Expert Systems with Applications*, [online] 143, p.112958. Available at: <<https://doi.org/10.1016/j.eswa.2019.112958>>.

MUSSINA, A., AUBAKIROV, S. AND TRIGO, P., 2018. Automatic Document Summarization based on Statistical Information. (Data), pp.71–76.

MUTLU, B., SEZER, E.A. AND AKCAYOL, M.A., 2020.

Candidate sentence selection for extractive text summarization. *Information Processing and Management*, [online] 57(6), p.102359. Available at: <<https://doi.org/10.1016/j.ipm.2020.102359>>.

NAJIBULLAH, A., 2015. Indonesian Text Summarization based on Naïve Bayes Method. *International Seminar and Conference 2015 : The Golden Triangle (Indonesia-India-Tiongkok)*, pp.67–78.

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.