

Pengujian Algoritma TextRank Dalam Merangkum Teks

Agustinus Yovi Siang Adi Setiawan*¹, Edwin Alexander²

¹Universitas Katolik Darma Cendika Surabaya

¹Surabaya, Indonesia

Email: 1agustinus.setiawan@student.ukdc.ac.id, 2edwin.alexander@ukdc.ac.id

Abstract

Summarizing text is key to overcoming issues with reading the entirety of a text. TextRank holds significant potential in summarizing texts. This research explores the utilization of the TextRank algorithm for automatic text summarization. The TextRank algorithm is a web approach that employs unsupervised methods and a graph-based ranking model. The research methodology encompasses qualitative methods, including data collection, data cleaning processes, TextRank algorithm application, and performance evaluation. TextRank can measure similarity between segments or sentences that contain similar content. In the application of the TextRank algorithm, the initial text summary of 111 words was reduced to 74 words.

Keywords: text summarization, textrank, data cleansing

Abstraksi

Peringkasan text merupakan kunci mengatasi permasalahan untuk membaca isi dari keseluruhan text. textRank berpotensi besar dalam merangkum teks. Penelitian ini mengeksplorasi penggunaan algoritma TextRank dalam melakukan ringkasan teks secara otomatis. Algoritma TextRank adalah pendekatan web yang menggunakan metode unsupervised dan model ranking berbasis grafik. Metode penelitian menggunakan metode kualitatif meliputi pengumpulan data, proses pembersihan data, penerapan algoritma TextRank, dan evaluasi kinerja. TextRank dapat mengukur kesamaan antar segmen atau kalimat yang memiliki konten serupa. Pada penerapan algoritma TextRank diperoleh hasil peringkasan teks yang awalnya berjumlah 111 kata menjadi 74 kata.

Kata Kunci: ringkasan teks, textrank, pembersihan data

1. PENDAHULUAN

Pada masa kini, informasi menjadi kebutuhan esensial yang tidak dapat diabaikan oleh siapa pun. Berbagai sumber menyajikan banyak informasi, salah satunya melalui media digital. Informasi yang diperoleh melalui media digital biasanya berupa artikel, blog, jurnal ilmiah, dan lain sebagainya [1]. Permasalahan yang sering didapatkan oleh para pembaca media digital adalah keterbatasan waktu untuk membaca keseluruhan isi dari dokumen teks yang dibaca guna mengetahui tujuan dan inti sari yang ada dalam dokumen. Oleh karena itu untuk mengenali informasi penting dari keseluruhan isi dokumen teks dan membuat ringkasan singkat dari dokumen yang teridentifikasi

tersebut diperlukan sistem yang dapat secara otomatis mengolah inti dari informasi dokumen teks [2]. Peringkasan teks otomatis yang akurat menjadi kunci untuk mengatasi permasalahan yang sering dihadapi oleh para pembaca.

Peringkasan teks otomatis merupakan penyajian suatu teks dalam bentuk yang lebih singkat dengan mengutip bagian penting dari satu ataupun lebih teks dan tidak lebih dari setengah teks asli [3]. Jumlah kalimat yang di ambil sebesar 20 hingga 50% dari keseluruhan kalimat yang terdapat dalam dokumen tersebut. Terdapat dua jenis tahap peringkasan teks otomatis yaitu peringkasan ekstraktif (*extractive summarization*) dan peringkasan abstraktif (*abstractive summarization*) [4]. Peringkasan ekstraktif merupakan proses peringkasan teks dengan mengekstrak informasi penting atau utama dari teks asli tanpa mengubah kata-kata atau struktur kalimat secara signifikan [5][6]. Sedangkan peringkasan teks abstraktif merupakan peringkasan teks dengan merekonstruksi informasi dalam kata atau kalimat baru yang merupakan inti sari informasi dari dokumen yang di ringkas [7]. Peneliti lain menjelaskan bahwa tujuan dari metode abstraktif ini adalah menciptakan kalimat singkat yang merangkum gagasan dari bacaan, mirip dengan cara manusia merangkum inti dokumen yang dibaca, dampaknya, kalimat yang diperoleh menjadi lebih sederhana, alami, dan fleksibel [8].

Ada beberapa algoritma yang dapat diterapkan dalam membuat peringkasan teks otomatis, salah satunya adalah algoritma *textrank*. *TextRank* adalah metode yang efektif dalam merangkum dokumen karena tidak membutuhkan data tambahan atau pengetahuan *spesifik dari dokumen yang hendak diringkas*[9]. *Textrank* merupakan metode berbasis web yang melakukan pendekatan *unsupervised* dan menggunakan pemodelan *graph-based ranking algorithm* (graf dengan model pemeringkatan) [10][11]. Peneliti lain juga menjelaskan bahwa *textrank* merupakan sebuah algoritma yang mengoperasikan sistem peringkat pada graf [12]. Pada penelitian ini bertujuan bertujuan untuk menerapkan algoritma *textrank* dalam merangkum teks.

2. METODE PENELITIAN

2.1. TextRank

TextRank merupakan algoritma yang digunakan untuk mengekstraksi kata kunci dan mempertimbangkan hubungan antar kata dalam teks berdasarkan graf [13]. *TextRank* dianggap sebagai metode ringkasan teks ekstraktif yang menggunakan model peringkat berbasis graf [10]. Metode ini menghasilkan graf terhubung sepenuhnya dan tidak berarah dari kalimat atau kata-kata dalam satu dokumen dan setiap kalimat dalam dokumen direpresentasikan oleh sebuah *node* [14]. *TextRank* telah menjadi metode yang banyak digunakan dalam penerapan untuk meringkas teks otomatis karena dapat mengekstrak pernyataan yang paling informatif dalam mengenali hubungan antar kalimat dalam sebuah dokumen. *TextRank* merupakan metode ringkasan yang dapat menghasilkan ringkasan yang lebih baik dengan menggunakan model topik dua tingkat [15].

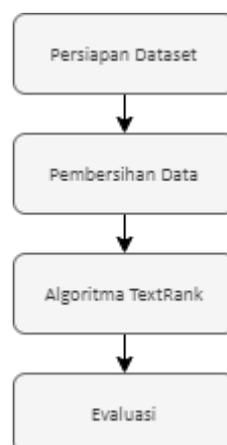
2.2. Text Summarizing

Text Summarizing (ringkasan teks) merujuk pada salah satu aplikasi pemrosesan bahasa alami (NLP) yang bertujuan untuk mengurangi teks panjang dan menguraikan poin-poin penting dan utama dalam teks sehingga membuat ringkasan yang koheren dan lebih singkat [16]. Ringkasan teks dalam istilah lain mengekstrak informasi paling penting dari sebuah teks untuk menghasilkan versi singkat. Ringkasan teks dalam beberapa bentuk sudah ada sejak tahun 1950-an, ringkasan teks telah menghasilkan banyak karya dibidang ilmu perpustakaan, penggunaan analisis linguistik dan penggunaan statistik [17]. Ringkasan teks diklasifikasikan dalam dua jenis dokumen yaitu dokumen Tunggal dan Multi-dokumen. Ringkasan dokumen Tunggal adalah prosedur otomatis yang berfokus pada ekstraksi informasi penting yang berguna dari satu teks seperti artikel, riset, atau halaman web. Sebaliknya, ringkasan Multi-dokumen adalah prosedur otomatis yang berfokus pada ekstraksi informasi penting dari beberapa teks yang berbagi topik yang sama, seperti berbagai jenis riset tentang masalah yang sama [14]. Prosedur ringkasan teks dapat diklasifikasikan berdasarkan kriteria berikut.

- 1) Jumlah dokumen yang di proses secara bersamaan berdasarkan jenis dokumen yaitu dokumen Tunggal dan Multi dokumen [14].
- 2) Jenis kalimat yang digunakan : ringkasan ekstraktif, ringkasan abstraktif dan ringkasan hibrida [18]. Ringkasan ekstraktif secara selektif memasukkan beberapa kalimat dari teks asli ke dalam ringkasan tanpa memodifikasi struktur kalimat dan ringkasan abstraktif mempertahankan makna teks tanpa mempertahankan strukturnya [19].
- 3) Ringkasan berbasis umum dan berbasis pertanyaan : ringkasan berbasis umum mengambil kalimat tanpa pertimbangan terhadap pertanyaan atau hubungan dengan judul, sedangkan ringkasan berbasis pertanyaan dikembalikan karena pertanyaan dan ringkasan antara judul dan kalimat [20].

2.3. Tahapan metode penelitian

Metode yang diterapkan pada penelitian ini adalah metode kualitatif untuk menguji algoritma *textrank* seperti yang ada pada Gambar 1 berikut ini.

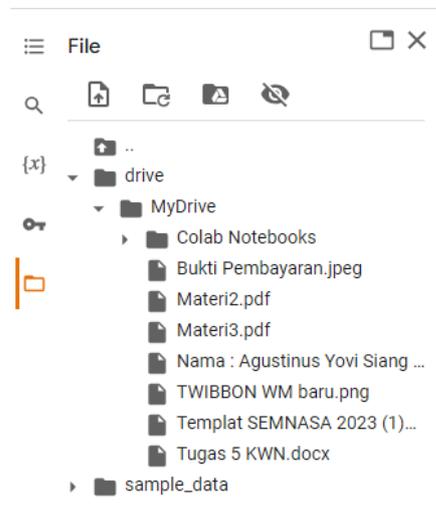


Gambar 1. Metode Pengujian Algoritma Textrank

Tahapan metode penelitian yang dilakukan dimulai dengan persiapan data set, kemudian melakukan pembersihan data, setelah pembersihan data selesai maka tahap selanjutnya pengaplikasian algoritma *textrank* kemudian yang terakhir melakukan evaluasi terhadap performa model. Sumber dataset berasal dari data sekunder yang diperoleh dari situs web yang dapat diperoleh pada pencarian internet.

2.3.1. Persiapan Dataset

Persiapan data set dilakukan dengan pengumpulan data berupa dokumen teks yang akan dilakukan pemrosesan menggunakan algoritma *textrank*. Data tersebut dapat ambil dari Google Drive yang telah terhubung dengan Google Colab. Data set yang digunakan adalah data dari teks berita yang diperoleh dari *website* Kompas.com.



Gambar 2. Data Google Drive terhubung dengan Google Colab

2.3.2. Pembersihan Data

Pembersihan data pada tahap ini meliputi pemisahan (*splitting*), tokenisasi (*tokenizing*), pemfilteran (*filtering*), penyeragaman karakter data (*case folding*), dan penghapusan kata yang tidak relevan (*stopward removal*).

2.3.3. Algoritma Textrank

Pada tahap ini, dilakukan inialisasi bobot awal dari sebuah kalimat dan dilakukan perhitungan kemiripan konten tumpang tindih untuk menghitung bobot hubungan antar kalimat. Setelah itu, dilakukan perhitungan *textrank* untuk mendapat bobot akhir dari setiap kalimat.

3. HASIL DAN PEMBAHASAN

Diberikan dokumen kata yang akan dianalisis. Contoh dokumen kata yang digunakan adalah artikel yang telah tayang di *Kompas.com* dengan judul “*Manchester United Ciptakan Rekor Buruk, Rp 289 Miliar untuk Pecat Ten Hag*”.

Manchester United perlu mengeluarkan biaya tinggi jika ingin memberhentikan sang pelatih, Eric Ten Hag. Kepemimpinan Eric ten Hag di *Manchester United* bergejolak menyusul hasil minor yang di dapat *Setan Merah* saat melawan *Newcastle* pada 16 besar Piala Liga Inggris 2023-2024. *Manchester United* besutan Eric ten Hag menelan kekalahan 0-3 dari *Newcastle* di stadion *Old Trafford*.

Ini menjadi kekalahan kedua secara beruntun *Manchester United* di stadion *Old Trafford*. Sebelumnya *Setan Merah* dibekuk *Manchester City* 0-3 dalam lanjutan *Premier League*. Menurut Squawka, *Manchester United* menuai rekor buruk karena kalah berturut-turut di kandang dengan skor 0-3 atau lebih untuk kali pertama dalam 61 tahun. Contoh dokumen yang diproses lebih lanjut pada pembersihan data sebagai berikut.

3.1. Melakukan *Splitting*, *Tokenizing*, *Filtering*, *Case Folding*, dan *Stopward Removal*

3.1.1. *Splitting*

Teks berita yang utuh akan dipisahkan menjadi beberapa kalimat. Pada proses ini beberapa teks dipisahkan atau dibagi menjadi beberapa kalimat dan sangat penting Karena dapat digunakan untuk struktur graf yang dapat dihitung peringkatnya dalam penentuan bagian penting dari teks tersebut. *Splitting* ditampilkan seperti dalam tabel 1 berikut.

Tabel 1. *Splitting*

No.	Kalimat
S1	Manchester United perlu mengeluarkan biaya tinggi jika ingin memberhentikan sang pelatih, Erik ten Hag.
S2	Kepemimpinan Erik ten Hag di Manchester United bergejolak menyusul hasil minor yang diperoleh Setan Merah saat melawan Newcastle pada 16 besar Piala Liga Inggris 2023-2024.
S3	Manchester United besutan Erik ten Hag menelan kekalahan 0-3 dari Newcastle di Stadion Old Trafford.
S4	Ini menjadi kekalahan kedua secara beruntun Manchester United di Stadion Old Trafford.
S5	Sebelumnya, Setan Merah dibekuk Manchester City 0-3 dalam lanjutan Premier League.
S6	Menurut Squawka, Manchester United menuai rekor buruk karena kalah berturut-turut di kandang dengan skor 0-3 atau lebih untuk kali pertama dalam 61 tahun.

3.1.2. *Tokenizing*

Pemotongan kalimat menjadi kata-kata tunggal dengan (" ") spasi sebagai *delimiter*. Dokumen teks ini dibagi menjadi token-token yang memisahkan setiap kata atau frasa menjadi simpul (*node*) dalam graf. *Tokenizing* ditampilkan dalam tabel 2 berikut.

Tabel 2. Tokenizing

No	Hasil Tokenizing		
S1	Manchester United perlu mengeluarkan biaya tinggi	Jika ingin memberhentikan sang pelatih	Erik ten Hag.
S2	Kepemimpinan Erik ten Hag di Manchester United bergejolak	menyusul hasil minor yang diperoleh Setan Merah Saat melawan	Newcastle dalam 16 besar Piala Liga Inggris 2023-2024.
S3	Manchester United besutan Erik ten	Hag menelan kekalahan 0-3 Dari Newcastle	Di Stadion Old Trafford.
S4	Ini menjadi kekalahan kedua	secara beruntun Man United	di Stadion Old Trafford.
S5	Sebelumnya, Setan Merah dibekuk	Manchester City 0-3 dalam	lanjutan Premier League.
S6	Menurut Squawka, Manchester United menuai rekor buruk karena	kalah berturut-turut di kandang dengan skor 0-3	atau lebih untuk kali pertama dalam 61 tahun.

3.1.3. Filtering

Pada dokumen teks ini, akan terjadi proses penghapusan tanda baca dan simbol yang tidak relevan, atau hanya mempertahankan entitas tertentu seperti kata benda atau kata kunci. *Filtering* ditampilkan dalam tabel 3 berikut :

Tabel 3. Filtering

No	Hasil Tokenizing		
S1	Manchester United perlu mengeluarkan biaya tinggi	Jika ingin memberhentikan sang pelatih	Erik ten Hag
S2	Kepemimpinan Erik ten Hag di Manchester	menyusul hasil minor yang diperoleh Setan Merah Saat	Newcastle dalam 16 besar Piala Liga Inggris 2023-2024

No	Hasil Tokenizing		
	United bergejolak	melawan	
S3	Manchester United besutan Erik ten	Hag menelan kekalahan 0-3 Dari Newcastle	Di Stadion Old Trafford
S4	Ini menjadi kekalahan kedua	secara beruntun Manchester United di	Stadion Old Trafford
S5	Sebelumnya, Setan Merah dibekuk	Manchester City 0-3 dalam lanjutan	Premier League
S6	Menurut Squawka, Manchester United menuai rekor buruk karena	kalah berturut-turut di kandang dengan skor 0-3	atau lebih untuk kali pertama dalam 61 tahun.

3.1.4. Case Folding

Pada proses ini, tahap yang dilakukan pada dokumen teks adalah mengubah seluruh karakter huruf dalam teks menjadi huruf kecil untuk memastikan keseragaman dalam representasi teks. *Case folding* pada proses ini membantu dalam meminimalkan ambiguitas yang muncul karena perbedaan huruf sehingga dapat memperbaiki konsistensi dalam perhitungan peringkat atau analisis yang dilakukan pada teks. *Case folding* ditampilkan dalam tabel 4 berikut ini

Tabel 4. Case Folding

No	Hasil Tokenizing		
S1	manchester united harus mengeluarkan biaya tinggi	jika ingin memberhentikan sang pelatih	erik ten hag
S2	kepemimpinan erik ten hag di manchester united bergejolak	menyusul hasil minor yang diperoleh setan merah saat melawan	newcastle dalam 16 besar piala liga inggris 2023-2024
S3	manchester	hag	di

No	Hasil Tokenizing		
	united besutan erik ten	menelan kekalahan 0-3 dari newcastle	stadion old trafford
S4	ini menjadi kekalahan kedua	secara beruntun manchester united di	stadion old trafford
S5	sebelumnya, setan merah dibekuk	man city 0-3 dalam lanjutan	premier league
S6	menurut squawka, manchester united menuai rekor buruk karena	kalah berturut- turut di kandang dengan skor 0-3	atau lebih untuk kali pertama dalam 61 tahun.

3.1.5. Stopword Removal

Stopword Removal merupakan tahap penghapusan kata-kata umum yang sering muncul. Kata-kata tersebut dihilangkan karena tidak memiliki makna [21]. Langkah proses *stopword removal* pada dokumen teks ini berguna untuk meningkatkan akurasi serta relevansi hasil dari algoritma *textrank* dengan menonjolkan informasi yang lebih penting dan relevan. *Stopword removal* ditampilkan dalam tabel 5 berikut:

Tabel 5. Stopword Removal

No.	Hasil Removal	Panjang kalimat
S1	Manchester United mengeluarkan biaya tinggi jika memberhentikan sang pelatih erik ten hag	12
S2	Kepemimpinan erik ten hag man united bergejolak hasil minor diperoleh setan merah saat melawan Newcastle 16 besar piala liga inggris 2023-2024	21
S3	Man united erik ten hag menelan kekalahan 0-3 newcastle stadion old Trafford	11
S4	Kekalahan kedua beruntun man united stadion old Trafford	8
S5	Setan merah dibekuk man city 0-3 lanjutan premier league	9
S6	Squawka man united menuai rekor buruk kalah berturut kendarang skor 0-3 lebih kali pertama 61 tahun	16

3.2. Algoritma Textrank

Pada tahap ini menghitung kemiripan antara segmen atau kalimat dengan konten yang tumpang tindih dengan menganggap kalimat yang sama dalam setiap segmen kalimat. Jika ada dua kata yang berada bersama-sama dalam satu kalimat maka dihitung sebagai 1. Berikut rumus kesamaan tumpang tindih konten pada rumus 1 berikut.

$$\text{Similarity}(S_i, S_j) = \frac{|\{W_k \mid W_k \in S_i \ \& \ W_k \in S_j\}|}{\log(|S_i|) + \log(|S_j|)} \quad (1)$$

Kemudian terapkan rumus kesamaan tumpang tindih konten pada semua kalimat. Sebagai contoh perhitungan kesamaan tumpang tindih konten untuk kalimat S1 dan S2 yaitu: $\text{Sim}(S_1, S_2) = 4 / (\log 12 + \log 21) = 1.6656$. Tabel 6 Berikut ini adalah hasil dari *similarity* terhadap persamaan tumpang tindih antar kalimat :

Tabel 6. *Similarity*

	S1	S2	S3	S4	S5	S6
S1	-	1.6656	1.8862	0.5044	0	0.4379
S2	1.6656	-	2.5384	0.8987	0.8785	0.7916
S3	1.8862	2.5384	-	2.0571	0.5010	0.8906
S4	0.5044	0.8987	2.0571	-	0.5384	0.9491
S5	0	0.8785	0.5010	0.5384	-	0.4633
S6	0.4379	0.7916	0.8906	0.9491	0.4633	-

3.3. Evaluasi

Berikut ini adalah hasil dari *text summarizing* dengan menggunakan algoritma TextRank: “Manchester United Ciptakan Rekor Buruk, Rp 289 Miliar untuk Pecat Ten Hag”. Manchester United harus mengeluarkan biaya tinggi seandainya mau memecat sang pelatih, Eric Ten Hag. Kepemimpinan Eric ten Hag di Manchester United bergejolak menyusul hasil minor yang diperoleh Setan Merah saat melawan Newcastle dalam 16 besar Piala Liga Inggris 2023-2024.

Menurut Squawka, Manchester United menuai rekor buruk karena kalah berturut-turut di kandang dengan skor 0-3 atau lebih untuk kali pertama dalam 61 tahun.

```

04 [25] print("Rangkuman:")
      print(summary)

```

Rangkuman:
 "Manchester United Ciptakan Rekor Buruk, Rp 289 Miliar untuk Pecat Ten Hag".Manchester United harus mengeluarkan biaya tinggi seandainya mau memecat sang pelatih, Menurut Squawka, Manchester United menuai rekor buruk karena kalah berturut-turut di kandang dengan skor 0-3 atau lebih untuk kali pertama dalam 61 tahun.

Gambar 3. Hasil text summarization menggunakan algoritma TextRank

Berdasarkan hasil dari pemisahan (*splitting*), tokenisasi (*tokenizing*), pemfilteran (*filtering*), penyeragaman karakter data (*case folding*), dan penghapusan kata yang tidak relevan (*stopward removal*) ringkasan yang diperoleh menggunakan algoritma TextRank dapat menunjukkan bahwa pada proses peringkasan dari jumlah kata awal sebanyak 111 kata mampu diringkaskan hingga 74 kata.

4. KESIMPULAN

Pengujian algoritma TextRank dalam merangkum teks menegaskan kemampuannya dalam menyusun ringkasan dengan mempertimbangkan hubungan dan bobot kata-kata dalam dokumen. Dalam penelitian ini, TextRank terbukti mampu mengidentifikasi kalimat-kalimat kunci yang relevan dengan inti dokumen, menghasilkan ringkasan yang informatif dengan memprioritaskan informasi penting. Hasil peringkasan dengan menggunakan algoritma TextRank menunjukkan bahwa proses peringkasan teks dari jumlah kata awal sebanyak 111 kata dapat dirangkum menjadi 74 kata. Meskipun demikian, keakuratan dan kebermanfaatannya sangat tergantung pada pengaturan parameter dan kemampuan untuk memahami konteks teks yang diberikan. Dengan demikian, TextRank menonjol sebagai algoritma potensial dalam merangkum teks, meskipun masih memerlukan pengembangan lebih lanjut guna meningkatkan kualitas rangkumannya. Metode penelitian yang mencakup langkah-langkah seperti persiapan dataset, pembersihan data, penerapan algoritma TextRank, serta evaluasi performa model turut mendukung pemahaman yang lebih baik terhadap kemampuan dan batasan algoritma ini dalam merangkum teks.

5. SARAN

Adapun saran yang dikembangkan untuk penelitian jurnal berjudul “Pengujian Algoritma TextRank dalam Merangkum Teks” adalah sebagai berikut :

1. Melakukan studi lebih mendalam untuk membandingkan kinerja algoritma TextRank dengan metode otomatis lain dalam merangkum teks.
2. Perluas data set untuk menguji TextRank pada berbagai jenis teks dan topik.
3. Teliti kemungkinan mengembangkan atau memodifikasi algoritma TextRank guna meningkatkan keakuratan dan efisiensi dalam merangkum teks secara otomatis.

DAFTAR PUSTAKA

- [1] M. Danuri, M. Informatika, J. Teknologi, and C. Semarang, “PERKEMBANGAN DAN TRANSFORMASI TEKNOLOGI DIGITAL.”
- [2] L. Al Qassem, D. Wang, H. Barada, A. Al-Rubaie, and N. Almoosa, “Automatic Arabic text summarization based on fuzzy logic,” in Proceedings of the 3rd international conference on natural language and speech processing, 2019, pp. 42–48.
- [3] M. F. Hidayattullah and A. Azizi, “Peringkasan Otomatis Teks Berbahasa Arab Menggunakan Algoritma TextRank,” *J. Ilm. Inform.*, vol. 6, no. 1, pp. 33–42, Jun. 2021, doi: 10.35316/jimi.v6i1.1231.
- [4] Y. Yuliska and K. U. Syaliman, “Literatur Review Terhadap Metode, Aplikasi dan Dataset Peringkasan Dokumen Teks Otomatis untuk Teks Berbahasa Indonesia,” *IT J. Res. Dev.*, vol. 5, no. 1, pp. 19–31, Jul. 2020, doi: 10.25299/itjrd.2020.vol5(1).4688.
- [5] D. P. Ismi and F. Ardianto, “Peringkasan Ekstraktif Teks Bahasa Indonesia dengan Pendekatan Unsupervised Menggunakan Metode Clustering.” *CYBERNETICS*, 2019.

- [6] D. A. Ridwansyah, "Peringkasan Teks Otomatis Ekstraktif Menggunakan Metode Non-Negative Matrix Facotrization Dan Word Sense Disambiguation Pada Berita Berbahasa Indonesia." Universitas Komputer Indonesia, 2021.
- [7] R. Bonovan, "Implementasi Pendekatan Abstraktif Pada Peringkasan Teks Bahasa Indonesia Menggunakan Metode Recurrent Neural Network." Universitas Komputer Indonesia, 2021.
- [8] K. Ivanedra and M. Mustikasari, "Implementasi Metode Recurrent Neural Network Pada Text Summarization Dengan Teknik Abstraktif," *J. Teknol. Inf. dan Ilmu Komput*, vol. 6, no. 4, p. 377, 2019.
- [9] L. Pertiwi, "Penerapan Algoritma Text Mining, Steaming Dan Texrank Dalam Peringkasan Bahasa Inggris," 2022.
- [10] R. Mihalcea and P. Tarau, "TextRank: Bringing Order into Texts."
- [11] M. A. Zamzam, "Sistem Automatic Text Summarization Menggunakan Algoritma Texrank." Universitas Islam Negeri Maulana Malik Ibrahim, 2019.
- [12] M. F. Fakhrezi, M. A. Bijaksana, and A. F. Huda, "Implementation of automatic text summarization with TextRank method in the development of Al-qur'an vocabulary encyclopedia," *Procedia Comput. Sci.*, vol. 179, pp. 391–398, 2021.
- [13] Y. Wang, H. Yin, and M. He, "Improvement of TextRank Based on Co-occurrence Word Pairs and Context Information BT - Smart Computing and Communication," 2018, pp. 226–235.
- [14] G. Amoudi, A. Almansour, and H. S. Alghamdi, "Improved Graph-Based Arabic Hotel Review Summarization Using Polarity Classification," *Appl. Sci.*, vol. 12, no. 21, p. 10980, 2022.
- [15] N. Akhtar, M. M. S. Beg, and H. Javed, "TextRank enhanced Topic Model for Query focussed Text Summarization," in *2019 Twelfth International Conference on Contemporary Computing (IC3)*, 2019, pp. 1–6, doi: 10.1109/IC3.2019.8844939.
- [16] Ö. Sahin, "Text Summarization BT - Develop Intelligent iOS Apps with Swift: Understand Texts, Classify Sentiments, and Autodetect Answers in Text Using NLP," Ö. Sahin, Ed. Berkeley, CA: Apress, 2021, pp. 121–135.
- [17] D. Shen, "Text Summarization BT - Encyclopedia of Database Systems," L. Liu and M. T. Özsu, Eds. New York, NY: Springer New York, 2018, pp. 4113–4117.
- [18] A. Elsaid, A. Mohammed, L. F. Ibrahim, and M. M. Sakre, "A comprehensive review of arabic text summarization," *IEEE Access*, vol. 10, pp. 38012–38030, 2022.
- [19] Z. Alami Merrouni, B. Frikh, and B. Ouhbi, "EXABSUM: a new text summarization approach for generating extractive and abstractive summaries," *J. Big Data*, vol. 10, no. 1, p. 163, 2023.
- [20] N. Alami, N. En-nahnahi, S. A. Ouatik, and M. Meknassi, "Using unsupervised deep learning for automatic summarization of Arabic documents," *Arab. J. Sci. Eng.*, vol. 43, pp. 7803–7815, 2018.
- [21] A. E. Budiman and A. Widjaja, "Analisis Pengaruh Teks Preprocessing Terhadap Deteksi Plagiarisme Pada Dokumen Tugas Akhir," *J. Tek. Inform. Dan Sist. Inf.*, vol. 6, no. 3, 2020.