

# Implementasi Text Summarization Pada Review Aplikasi Digital Library System Menggunakan Metode *Maximum Marginal Relevance*

Ichwanul Muslim Karo Karo\*<sup>1</sup>, Adidtya Perdana<sup>2</sup>, Sri Dewi<sup>3</sup>

<sup>1,2,3</sup>Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam/Universitas Negeri Medan

e-mail: \*[ichwanul@unimed.ac.id](mailto:ichwanul@unimed.ac.id), [adidtya@unimed.ac.id](mailto:adidtya@unimed.ac.id), [sridewi@unimed.ac.id](mailto:sridewi@unimed.ac.id)

**Abstract** – Text summarization refers to the creation of text summaries automatically with the natural language processing approach (NLP). Text summarization is required when the number of documents or reviews to be summarized is large. A resulting summary can be knowledge, input, or advice for the improvement or development of various applications. The Digital Library System Application is a mobile app for State University of Medan library services (Unimed). The app has a lot of reviews on various platforms. Of course, the review summary is a user experience and can be an input to developing the latest version. However, it would be a challenge if the entire user review were summarized manually, as it would take a long time. The research aims to provide a summary of the mobile app reviews with an automatic text synthesis approach. The algorithms used in the summary of the text in this study are Maximum Marginal Relevance (MMR) and evaluation processes using precision, recall, and F1. The review will go through the text pre-processing phase with the help of the NLTK library. The study successfully identified 30 reviews with the highest MMR score. Furthermore, the review summary presented is a series of 10 reviews with the highest MMR rating. The resulting summary has an accuracy of 30.51%, a recall of 56.25%, and an F1 score of 39.56%.

**Abstrak** – Peringkasan teks merujuk pada pembuatan rangkuman teks secara otomatis dengan pendekatan *natural language processing* (NLP). *Text summarization* dibutuhkan saat jumlah dokumen atau review yang akan dirangkum dalam jumlah yang banyak. Sebuah rangkuman yang dihasilkan dapat menjadi pengetahuan, masukan maupun saran untuk perbaikan/pengembangan berbagai aplikasi. Aplikasi Digital Library System merupakan sebuah mobile apps untuk layanan perpustakaan Universitas Negeri Medan (Unimed). Aplikasi tersebut memiliki banyak ulasan di berbagai platform. Tentu, rangkuman ulasan tersebut merupakan pengalaman pengguna dan dapat menjadi masukan untuk pengembangan versi terbaru. Namun menjadi tantangan jika seluruh ulasan pengguna dirangkum secara manual, karena akan memakan waktu yang lama. Penelitian ini bertujuan untuk menyediakan rangkuman atas ulasan mobile Apps tersebut dengan pendekatan peringkasan teks secara otomatis. Algoritma yang digunakan dalam peringkasan teks di penelitian ini ialah Maximum Marginal Relevance (MMR) dan proses evaluasi menggunakan presisi, recall dan F1. Ulasan mobile apps diperoleh dari play store dan App Store. Ulasan akan melalui tahapan *text pre-processing* dengan bantuan *library NLTK*. Penelitian ini berhasil mengidentifikasi 30 review dengan nilai MMR tertinggi. Lebih lanjut, rangkuman ulasan yang disajikan merupakan rangkaian 10 ulasan dengan nilai MMR tertinggi. Rangkuman yang dihasilkan memiliki tingkat presisi sebesar 30.51%, recall sebesar 56.25%, dan skor F1 sebesar 39.56%.

**Kata Kunci** – Peringkasan teks, Mobile Apps, NLP, MMR, Presisi, Recall, F1

## I. PENDAHULUAN

Peringkasan teks (*text summarization*) adalah pembuatan rangkuman teks secara otomatis melibatkan algoritma [1] dan menggunakan komputer [2]. Prinsip kerjanya dengan mengurangi dokumen teks dan menciptakan ringkasan yang tetap mempertahankan poin penting dari dokumen asli [3]. Sebuah sistem peringkasan diberi masukan berupa teks, kemudian melakukan peringkasan, dan menghasilkan keluaran berupa teks yang lebih singkat dari teks aslinya. Pada perkembangannya, *text summarization* telah diimplementasikan untuk merangkum berbagai dokumen, artikel maupun ulasan. Text summarization dianggap mampu menjadi solusi dalam menyediakan rangkuman dengan cepat dibandingkan dengan peringkasan secara manual [4].

Metode *text summarization* dapat dibagi menjadi dua kategori utama: *extractive* dan *abstractive* [3], [5]. Pendekatan *extractive* memilih kalimat-kalimat utama dari teks sumber [6], [7], [8], sedangkan pendekatan *abstractive* menghasilkan ringkasan baru dengan menggunakan kata-kata dan struktur kalimat yang mungkin tidak ada dalam teks asli [9], [10]. Berdasarkan penelitian oleh [11], [12], pendekatan ekstraksi lebih sederhana dan lebih mudah diimplementasikan dalam berbagai studi kasus. Sehingga dalam penelitian ini menggunakan pendekatan ekstraksi pula.

Maximum Marginal Relevance (MMR) adalah salah satu teknik ekstraksi ringkasan yang diterapkan untuk merangkum dokumen tunggal atau sejumlah dokumen [13]. MMR adalah metode peringkasan yang bertujuan untuk

mengekstrak informasi yang relevan tanpa adanya redundansi. MMR merupakan salah satu metode ekstraksi ringkasan (*extractive summary*) yang digunakan untuk meringkas dokumen tunggal atau multi dokumen. Algoritma MMR melakukan peringkasan teks dengan menggabungkan matriks cosine similarity di antara setiap kalimat [14]. Terdapat dua jenis kemiripan yang dievaluasi, yaitu kemiripan antar dokumen atau kalimat, dan kemiripan antara dokumen atau kalimat dengan pertanyaan/query. Algoritma MMR merangkum dokumen dengan menghitung derajat kemiripan (similarity) antar bagian teks. MMR tertinggi dipilih secara iteratif dari resume sampai nilai yang telah ditentukan dan tercapai. Penelitian-penelitian sebelumnya juga telah banyak menggunakan metode MMR untuk berbagai kasus, seperti menyediakan rangkuman atas ulasan berbagai produk (misalnya *mobile apps*, barang di *e-commerce*) maupun layanan.

Dengan perkembangan penelitian dibidang *text summarization* yang cukup pesat, maka sangat memungkinkan untuk menyediakan rangkuman atas berbagai ulasan *mobile apps*. Penelitian ini bertujuan untuk mengimplementasikan algoritma MMR untuk menyajikan rangkuman dari ulasan aplikasi Digital Library System. Aplikasi Digital Library System adalah sebuah platform online berbasis aplikasi yang digunakan oleh pihak perpustakaan Universitas Negeri Medan (Unimed) untuk mempermudah pelayanan kepada mahasiswa. Aplikasi Digital Library System memberikan banyak fitur termasuk fitur peminjaman buku, pengembalian hingga scan barcode untuk memasuki ruangan di Digital Library Unimed [15]. Ulasan dari aplikasi tersebut diperoleh dari Google Play Store dan Apps Store.

Sejauh penelitian ini, aplikasi tersebut belum pernah tersedia versi terbaru dan menyediakan rangkuman atas ulasannya. Penelitian ini berkontribusi menghasilkan sebuah rangkuman yang dapat menjadi sebuah masukan untuk pengembangan versi terbaru dari aplikasi tersebut. Lebih lanjut, kualitas rangkuman juga di evaluasi kualitas metrik ROUGE (precision, recall dan F1).

## II. PENELITIAN YANG TERKAIT

Terdapat beberapa penelitian sebelumnya yang telah mengimplementasikan *text summarization* menggunakan metode MMR untuk berbagai kasus. Implementasi *text summarization* pada Review Aplikasi Super di Google Play Store menggunakan metode MMR [16]. Penelitian meringkas dokumen tunggal atau multi dokumen. Hasil dari penelitian yang menggunakan metode MMR menunjukkan nilai presisi dengan rata-rata sebesar 40,4% dalam tiga kali percobaan, dengan nilai presisi tertinggi mencapai 60,4% pada percobaan yang menggunakan parameter  $\lambda = 0,7$ . Meskipun memberikan hasil terbaik, namun pada dokumen dalam skala besar, semakin kecil nilai  $\lambda$  dapat mengurangi kemungkinan adanya hasil ringkasan yang bersifat redundan.

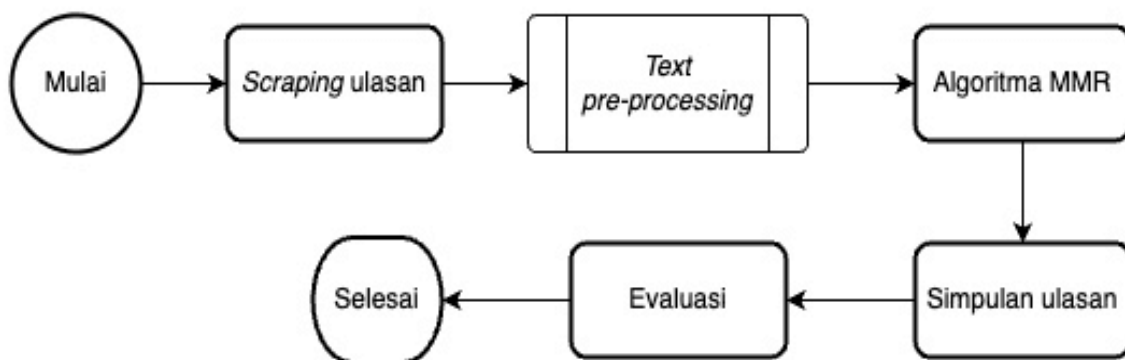
Selain untuk meringkas review, algoritma MMR juga telah digunakan untuk meringkan artikel maupun berita. Penelitian oleh [17] menggunakan algoritma MMR untuk meringkas artikel berita pendek. Hasil yang diperoleh untuk seluruh artikel berita pendek yang diuji pada 30 artikel, diperoleh dengan nilai akurasi 59.5 % metode cosine coefficient ini lebih unggul diterapkan untuk proses peringkasan artikel berita pendek dibandingkan dengan metode jaccard dan dice coefficient. Dengan demikian, terbukti bahwa peringkasan artikel berita pendek lebih tepat menggunakan metode MMR berbasis cosine coefficient. Penelitian lainnya oleh [18] menyajikan ringkasan artikel berita terkait COVID-19 menggunakan metode MMR. Penelitian ini dilakukan dengan mengambil 30 sampel data artikel berita dengan tema COVID-19 dari website penyedia berita online kompas.com. Diperoleh hasil pengujian dengan nilai, koefisien pengatur nilai terbaik adalah  $\alpha=0,5$  dengan hasil precision = 0,684333, recall = 0,772 dan f-measure = 0,7. Sedangkan berdasarkan jumlah kata, jumlah kata dibah 300 menghasilkan nilai f-measure terbaik dengan nilai 0,726923.

Sedangkan untuk artikel berita berbahasa Indonesia yang lain juga pernah dikembangkan peringkasan teks otomatis menggunakan algoritma MMR oleh [19]. Berdasarkan hasil pengujian peringkasan teks otomatis menggunakan MMR didapatkan hasil rata-rata precision sebesar 0,70, rata-rata recall 0,75, rata-rata f-measure adalah 0,70 dan rata-rata akurasi sebesar 74,17% dari 150 data yang diuji. Hasil ringkasan berhasil mendapatkan kalimat yang relevan dengan isi dari dokumen dan berhasil mengurangi redundant dari kalimat-kalimat yang sudah dijadikan ringkasan. Penerapan serupa juga pernah dilakukan oleh [14]. Hasil dari sistem peringkasan teks otomatis dalam penelitian ini berupa kalimat inti yang mirip dengan query dan berdasarkan urutan bobot, jadi untuk pengembangan penelitian berikutnya diharapkan hasil ringkasan memiliki urutan berdasarkan sistematika yang baik.

## III. METODE PENELITIAN

Adapun tahapan implementasi *text summarization* yang dilakukan terhadap data ulasan aplikasi unimed library system menggunakan metode MMR dapat dilihat Gbr.1. Tahapan penelitian dimulai dengan proses ekstraksi ulasan aplikasi unimed library system. Tahapan selanjutnya ialah praproses ulasan. Tahapan praproses meliputi penerjemahan kedalam bahasa Indonesia dan penghapusan simbol dan lainnya. Tahap intinya ialah pembuatan ringkasan referensi untuk setiap ulasan dengan algoritma MMR. Algoritma tersebut digunakan untuk memilih

kalimat-kalimat yang paling relevan dengan ringkasan yang telah dipilih sebelumnya dan juga dengan ringkasan referensi. Setiap kalimat dipilih berdasarkan nilai kemiripan tertinggi dengan mempertimbangkan aspek-aspek keterkaitan kalimat internal dan keterkaitan dengan referensi. Proses ini diulang untuk setiap ulasan hingga tercapai jumlah kalimat yang diinginkan untuk setiap ringkasan. Selanjutnya, dilakukan evaluasi menggunakan ROUGE untuk mengukur kualitas ringkasan yang dihasilkan dengan membandingkannya dengan konten asli [7], [20]. Hasil evaluasi, yaitu presisi, recall, dan F1 score, ditampilkan atau dicetak untuk memberikan pemahaman tentang kinerja metode MMR tersebut.



Gbr. 1 Alur penelitian

#### A. Dataset Review

Data yang menjadi fokus penelitian ini diperoleh melalui telaah mendalam terhadap ulasan dan *feedback* pengguna pada aplikasi sistem perpustakaan digital yang tersedia di platform distribusi Google Playstore dan Apps Store. Proses pengambilan ulasan tersebut menggunakan teknik *scraping*. *Scraping* data adalah proses mengambil atau mengumpulkan data dari internet. Data yang diambil dapat berupa teks, gambar, video, dll. Proses ini dilakukan dengan menggunakan perangkat lunak khusus yang disebut *web scraper* [21], [22].

Informasi yang digunakan dalam penelitian ini diambil dari ulasan tersebut merangkum pandangan dan pengalaman pengguna terhadap aplikasi-aplikasi tersebut, mencakup aspek-aspek seperti kegunaan, antarmuka pengguna, performa, dan fitur-fitur yang disediakan. Data yang digunakan dalam penelitian ini secara substansial bersifat teks, mencakup beragam ekspresi bahasa yang digunakan oleh pengguna dalam merinci pengalaman mereka dengan aplikasi-aplikasi perpustakaan digital.

#### B. Text Preprocessing

*Text preprocessing* memainkan peran penting dalam peringkasan teks [23] bahkan dalam beberapa kasus mampu meningkatkan performansi [24], [25]. *Text preprocessing* adalah proses mengubah data yang tidak terstruktur menjadi data yang terstruktur dengan menggunakan pemrosesan bahasa alami (NLP) [26]. Dengan kata lain, tahap ini mengubah data tekstual *input* menjadi format yang bersih dan konsisten yang dapat diumpungkan ke model peringkasan yang diusulkan. Penelitian ini menerapkan beberapa teknik pra-pemrosesan teks seperti penghilangan *stopword*, tokenisasi, *case folding*, penghilangan simbol, penghilangan emoji, dan *lemmatization*. Untuk implementasi fase-fase pra-pemrosesan tersebut, penelitian ini menggunakan paket Python NLTK seperti yang telah digunakan pada penelitian-penelitian sebelumnya [7], [27], [28].

#### C. Metode Maximum Marginal Relevance

Algoritma (MMR) merupakan salah satu metode ekstraksi ringkasan (extractive summary) [29]. Algoritma tersebut mampu merangkum dokumen tunggal atau multi dokumen dengan mempertimbangkan kesamaan antar bagian teks [29], [30]. Dalam proses peringkasan dokumen menggunakan MMR, dokumen disegmentasi menjadi kalimat, dan kemudian dilakukan pengelompokan berdasarkan keterkaitan kalimat tersebut. MMR mengaplikasikan matriks cosine similarity untuk merangking kalimat-kalimat sebagai respons terhadap query yang diberikan oleh [13]. Cara dari algoritma MMR untuk meringkas dokumen adalah dengan menghitung kesamaan antar bagian teks [18]. MMR adalah teknik peringkasan yang memiliki tujuan untuk mengambil informasi yang relevan dan tidak mengandung redundansi [19]. Penelitian ini merujuk kerja sebagaimana penelitian sebelumnya [14].



### C. Ringkasan Ulasan

Bagian menyajikan ringkasan ulasan yang telah di proses menggunakan algoritma MMR. Dari 461 ulasan, diiterasi pertama diperoleh top 30 ulasan dengan nilai MMR tertinggi. Selanjutnya, algoritma memproses di iterasi kedua. Luaran iterasi kedua, diperoleh top 10 ulasan (Gbr. 3). Dari ulasan tersebut diidentifikasi topik yang dapat menjadi sebuah masukan maupun *insight* untuk pengembangan aplikasi.

Berdasarkan simpulan yang diperoleh, terdapat 2 topik utama yang menjadi fokus ulasan. Ulasan banyak membahas fitur dan Wi-fi dari aplikasi tersebut. Fitur yang ulasan meliputi ketersediaan yang masih sedikit dan perlu ditingkatkan. Mobile apps sering terkedala login, padahal aksesny tidak terlalu buruk.

```
0  aplikasinya bagus fitur yg sedikitsekedar sara...
1  bagus sih sangat disayangkan penarikan kartu f...
2  wifinya lemot putus jaringan semoga ditingkatkan
3  kartu anggota fisik tarik pembuatnya bayar dis...
4  aplikasi buruk sua buku keterangan nya berbali...
5  aduhh aplikasi ribet yaa aplikasi memperpanjan...
6  ga sesuai gangan ekspektasi minggunya peratura...
7  halo mahasiswa aplikasi aplikasi membantu
8  aplikasi nya mantap update sewingga memudahkan...
9  mantap tampilannya bagusklo fitur profil foto ...
Name: summary, dtype: object
```

Gbr. 3 Top 10 ulasan berdasarkan nilai MMR

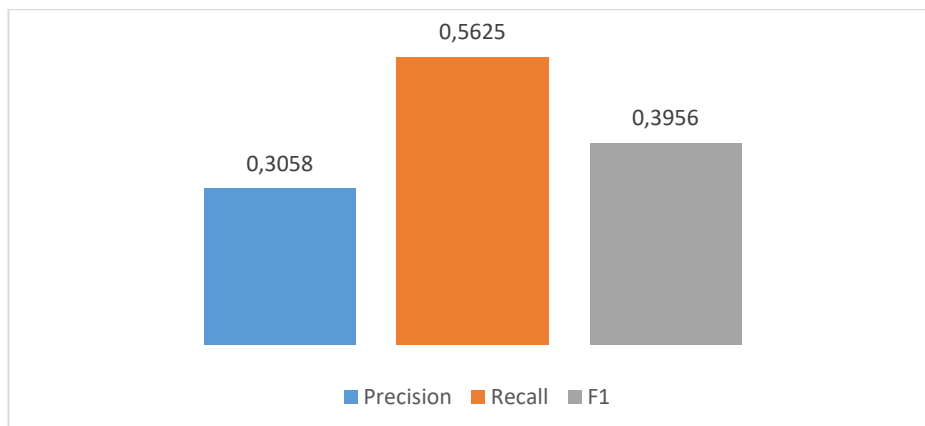
### D. Evaluasi

Tahapan akhir ialah mengevaluasi ringkasan ulasan yang dihasilkan. Proses evaluasi dengan menggunakan ROUGE meliputi presisi, recall, dan skor F. Presisi dalam konteks ROUGE mencerminkan fraksi  $n$ -gram dalam simpulan yang dihasilkan yang juga ada dalam referensi. Adapun formula untuk menghitung precision dengan menggunakan persamaan (1). Recall adalah fraksi  $n$ -gram referensi yang juga muncul dalam simpulan yang dihasilkan oleh algoritma. Adapun formula untuk menghitung recall dengan menggunakan persamaan (2). F1 adalah kombinasi atas kedua metric precision dan recall (persamaan (3)).

$$RAUGE_{precision} = \frac{unigram\ simpulan \cap unigram\ referensi}{|unigram\ simpulan|} \quad (1)$$

$$RAUGE_{recall} = \frac{unigram\ simpulan \cap unigram\ referensi}{|unigram\ referensi|} \quad (2)$$

$$RAUGE_{F1} = \frac{2 \cdot recall \cdot precision}{recall + precision} \quad (3)$$



Gbr. 4 Performansi MMR dalam menyajikan ringkasan ulasan

Performansi algoritma MMR dalam menyajikan simpulan dapat dilihat pada Gbr.4. Kualitas simpulan atas ulasan aplikasi digital library system yang dihasilkan diukur dengan dengan metric precision, recall dan F1. Berdasarkan gambar tersebut, tingkat presisi simpulan yang dihasilkan ialah 30,58%, kondisi ini menandakan tingkat presisi simpulan yang dihasilkan masih cukup rendah. Rating recall yang dihasilkan ialah 56,25%, interpretasinya ialah simpulan yang dihasilkan sudah cukup baik. Sedangkan nilai F1 nya adalah 39,56% juga menandakan masih rendah. Secara umum kualitas simpulan yang dihasilkan belumlah cukup optimal, hal ini ditandai dengan performansi atas masing masing metric yang masih dibawah 65%.

## V. KESIMPULAN

Ketersediaan ulasan atas aplikasi digital librabry system yang sangat banyak, sehingga membutuhkan text summarization dalam menyediakan simpulan dengan cepat. Tentunya hal tersebut sejalan dengan tujuan penelitian ini. Algoritma MMR sebagai salah satu algoritma text summarization dapat diimplementasikan untuk menyediakan ringkasan atas seluruh ulasan yang diperoleh. Namun performansi algoritma MMR yang dihasilkan memiliki tingkat presisi sebesar 30.51%, recall sebesar 56.25%, dan skor F1 sebesar 39.56%. Meskipun tingkat recall cukup baik, tingkat presisi yang relatif rendah menunjukkan bahwa sebagian besar kalimat yang dipilih sebagai ringkasan mungkin termasuk beberapa kalimat yang kurang relevan. Dengan nilai F1 yang mencerminkan keseimbangan antara presisi dan recall, masih perlu diperhatikan dan dioptimalkan untuk meningkatkan kualitas ringkasan yang dihasilkan oleh metode MMR. Evaluasi ini memberikan pandangan tentang kinerja sistem dalam menangkap informasi yang relevan dari ulasan, namun perlu diperhatikan bahwa hasil ini dapat dipengaruhi oleh sejumlah faktor, termasuk pengaturan parameter metode MMR dan kualitas referensi ringkasan yang digunakan dalam evaluasi.

## UCAPAN TERIMA KASIH

Ucapan terima kasih penulis kepada Universitas Negeri Medan melalui anggaran PNPB Unimed tahun 2024 dan Big Data & AI research center atas dukungan dukungan terkait dana penelitian dan bantuan fasilitas penelitian.

## DAFTAR PUSTAKA

- [1] M. A. Zamzam, "SISTEM AUTOMATIC TEXT SUMMARIZATION MENGGUNAKAN ALGORITMA TEXTRANK," *MATICS*, vol. 12, no. 2, 2020, doi: 10.18860/mat.v12i2.8372.
- [2] A. P. Widyassari *et al.*, "Review of automatic text summarization techniques & methods," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 4. 2022. doi: 10.1016/j.jksuci.2020.05.006.
- [3] W. S. El-Kassas, C. R. Salama, A. A. Rafea, and H. K. Mohamed, "Automatic text summarization: A comprehensive survey," *Expert Systems with Applications*, vol. 165. 2021. doi: 10.1016/j.eswa.2020.113679.
- [4] N. K. Nagwani, "Summarizing large text collection using topic modeling and clustering based on MapReduce framework," *J Big Data*, vol. 2, no. 1, 2015, doi: 10.1186/s40537-015-0020-5.
- [5] A. D. Dhawale, S. B. Kulkarni, and V. M. Kumbhakarna, "Experimental Evaluation and Approach of Enhancement in Generation of Automatic Unsupervised Extractive Text Summarization of Marathi Text By Using Machine Learning Algorithm," *Journal of Machine and Computing*, vol. 2, no. 1, 2022, doi: 10.53759/7669/jmc202202004.
- [6] Y. K. Khor, C. W. Tan, and T. M. Lim, "Extractive Summarization on Food Reviews," *The Journal of The Institution of Engineers, Malaysia*, vol. 82, no. 3, 2022, doi: 10.54552/v82i3.96.
- [7] V. Gulati, D. Kumar, D. E. Popescu, and J. D. Hemanth, "Extractive Article Summarization Using Integrated TextRank and BM25+ Algorithm," *Electronics (Switzerland)*, vol. 12, no. 2, 2023, doi: 10.3390/electronics12020372.
- [8] K. Ramani, K. Bhavana, A. Akshaya, K. S. Harshita, C. R. Thoran Kumar, and M. Srikanth, "An Explorative Study on Extractive Text Summarization through k-means, LSA, and TextRank," in *WiSPNET 2023 - International Conference on Wireless Communications, Signal Processing and Networking*, 2023. doi: 10.1109/WiSPNET57748.2023.10134303.
- [9] A. A. Syed, F. L. Gaol, and T. Matsuo, "A survey of the state-of-the-art models in neural abstractive text summarization," *IEEE Access*, vol. 9. 2021. doi: 10.1109/ACCESS.2021.3052783.
- [10] S. Gupta and S. K. Gupta, "Abstractive summarization: An overview of the state of the art," *Expert Systems with Applications*, vol. 121. 2019. doi: 10.1016/j.eswa.2018.12.011.
- [11] C. Fang, D. Mu, Z. Deng, and Z. Wu, "Word-sentence co-ranking for automatic extractive text summarization," *Expert Syst Appl*, vol. 72, 2017, doi: 10.1016/j.eswa.2016.12.021.
- [12] A. K. Yadav, M. Kumar, and A. Pathre, "Implemented Text Rank based Automatic Text Summarization using Keyword Extraction," *International Research Journal of Innovations in Engineering and Technology*, vol. 04, no. 11, 2020, doi: 10.47001/irjiet/2020.411003.
- [13] R. Robiyanto, N. Nugraha, and I. Apriatna, "Peringkasan Teks Otomatis Berita Menggunakan Metode Maximum Marginal Relevance," *JEJARING: Jurnal Teknologi dan Manajemen Informatika*, vol. 4, no. 1, 2019, doi: 10.25134/jejaring.v4i1.6712.
- [14] A. Kurniawan and Mohd. I. Humaidy, "Penerapan Algoritma Maximum Marginal Relevance Dalam Peringkasan Teks Secara Otomatis," *Bulletin of Data Science*, vol. 1, no. 2, 2022.

- [15] Y. Yusniah, R. S. Asri, P. A. Parent, and N. Nuraina, "Analisis Konsep Kerjasama Eksternal Antar Perpustakaan di Perguruan Tinggi," *Da'watuna: Journal of Communication and Islamic Broadcasting*, vol. 3, no. 2, 2022, doi: 10.47467/dawatuna.v3i2.2467.
- [16] D. A. Louis, S. Rostianingsih, and L. W. Santoso, "Implementasi Text Summarization pada Review Aplikasi Super di Google Play Store Menggunakan Metode Maximum Marginal Relevance," *Jurnal Infra*, vol. 10, no. 2, 2022.
- [17] Arie Atwa Magriyanti, "MAXIMUM MARGINAL RELEVANCE BERBASIS BOOLEAN MODEL PADA PERINGKASAN ARTIKEL BERITA PENDEK," *Jurnal Ilmiah Teknik Informatika dan Komunikasi*, vol. 1, no. 3, 2021, doi: 10.55606/juitik.v1i3.132.
- [18] Y. Ananda Kresna, I. Cholissodin, and Indriati, "Peringkasan Teks Menggunakan Metode Maximum Marginal Relevance terhadap Artikel Berita terkait COVID-19," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 9, 2021.
- [19] N. F. Saraswati, Indriati, and R. S. Perdana, "Peringkasan Teks Otomatis Menggunakan Metode Maximum Marginal Relevance Pada Hasil Pencarian Sistem Temu Kembali Informasi Untuk Artikel Berbahasa Indonesia," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya*, vol. 2, no. 11, 2018, doi: 10.1016/s1010-6030(01)00380-x.
- [20] C. Y. Lin, "Rouge: A package for automatic evaluation of summaries," *Proceedings of the workshop on text summarization branches out (WAS 2004)*, no. 1, 2004.
- [21] D. Glez-Peña, A. Lourenço, H. López-Fernández, M. Reboiro-Jato, and F. Fdez-Riverola, "Web scraping technologies in an API world," *Brief Bioinform*, vol. 15, no. 5, 2013, doi: 10.1093/bib/bbt026.
- [22] M. A. Khder, "Web scraping or web crawling: State of art, techniques, approaches and application," *International Journal of Advances in Soft Computing and its Applications*, vol. 13, no. 3, 2021, doi: 10.15849/ijasca.211128.11.
- [23] S. Das, "ESStep: A novel method for semantic text summarization with web-based big data," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, 2019, doi: 10.35940/ijrte.C5802.098319.
- [24] B. Baykara and T. Güngör, "Turkish abstractive text summarization using pretrained sequence-to-sequence models," *Nat Lang Eng*, vol. 29, no. 5, 2023, doi: 10.1017/S1351324922000195.
- [25] Z. Rahimi and M. M. Homayounpour, "The impact of preprocessing on word embedding quality: a comparative study," *Lang Resour Eval*, vol. 57, no. 1, 2023, doi: 10.1007/s10579-022-09620-5.
- [26] I. M. Karo Karo, M. Farhan, M. Fudzee, S. Kasim, and A. A. Ramli, "Karonese Sentiment Analysis: A New Dataset and Preliminary Result," *JOIV: International Journal on Informatics Visualization*, vol. 6, no. 2–2, pp. 523–530, 2022, [Online]. Available: [www.joiv.org/index.php/joiv](http://www.joiv.org/index.php/joiv)
- [27] R. Yanuarti and H. A. Al Faruq, "Implementasi Text Summarization Pada Reading Comprehension Menggunakan Library Python," *Jurnal Aplikasi Sistem Informasi Dan Elektronika*, vol. 2, no. 1, 2022.
- [28] E. J. Rifano, Abd. C. Fauzan, A. Makhi, E. Nadya, Z. Nasikin, and F. N. Putra, "Text Summarization Menggunakan Library Natural Language Toolkit (NLTK) Berbasis Pemrograman Python," *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 2, no. 1, 2020, doi: 10.28926/ilkomnika.v2i1.32.
- [29] A. G. L. Babu and S. Badugu, "Extractive Summarization of Telugu Text Using Modified Text Rank and Maximum Marginal Relevance," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 22, no. 9, 2023, doi: 10.1145/3600224.
- [30] D. Delvin, D. Arisandi, and T. Sutrisno, "APLIKASI PERINGKASAN DOKUMEN MENGGUNAKAN METODE MAXIMUM MARGINAL RELEVANCE (MMR)," *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 10, no. 1, 2022, doi: 10.24912/jiksi.v10i1.17820.