COMPARISON OF MAXIMAL MARGINAL RELEVANCE (MMR) AND TEXTRANK AUTOMATIC TEXT SUMMARIZATION METHODS IN JOURNALS

Muhammad Alif Al Fattah[1] **Rizal**[2*] **Rini Meiyanti** [3]

[1]Information Engineering, Universitas Malikussaleh, Aceh, Lhokseumawe, 24353, Indonesia, Muhammad.alif.190170070@mhs.unimal.ac.id

[2]Information Engineering, Universitas Malikussaleh, Aceh, Lhokseumawe, 24353, Indonesia, rizal@unimal.ac.id

[3]Information Engineering, Universitas Malikussaleh, Aceh, Lhokseumawe, 24353, Indonesia, rinimeiyanti@unimal.ac.id

✉Corresponding Author: **rizal@unimal.ac.id |**

Abstract

The purpose of this study is to find out the application of the Maximal Marginal Relevance (MMR) and TextRank methods in automatic text summarization in journals in viewing the best model value in text automatically. This study can also implement an automatic text summarization application and find out the comparison between MMR and TextRank in the text summarization process in journals. Next research will evaluate the performance of the two models in producing relevant and informative text summaries. The problem of this research is how to overcome the problem of summarizing text with the basic concept of summary in providing the essence or overall content text in a journal. The main focus of this research is to display significant and relevant information in a more organized form and to display values for the efficiency of which model is the best after comparing the two models. The results of this research are a decision support system for determining the quality of poor rice using the fuzzy madm yager model with a value of (MMR) 0.4 and top N(3). The results of the comparison of the Maximal Marginal Relevance (MMR) similarity values are 0.510 and the score is 0.510, while the TextRank similarity is 0.510 and the score is 0.015. Based on testing of the two models, the best value was obtained from the rextrank model with a score of 0.015.

Keywords: maximal marginal relevance (mmr), textrank, summary

Introduction

Background

The rapid development of the academic world, research, scientific journals often have long and complex content in taking a scientific summary [1]. Understanding the many existing journals is a challenge for researchers who are members of research-related fields [2] [3]. The process of understanding a journal thoroughly requires a lot of time and effort, especially when researchers have to access many journals to get correct and relevant information in search journals. [4]. In today's digital era, information is easy to obtain and can be viewed quickly in journals or scientific works found in online media [5] [6]. However, in the face of obstacles in the form of long journals and the difficulty of obtaining a comprehensive and fast understanding of the topic being researched and the difficulty of processing information relating to related fields to draw conclusions [7].

A journal is the result of research that discusses a problem according to a field of science that has been studied by lecturers, practitioners or students [8]. The main focus of this research is to present better, relevant and organized information in looking at the results of the comparison of the two models and the summarization process is one of the most important things in the journal browsing process [9]. The basic concept of this research is to provide an important understanding of the overall content of the article/journal which will provide concise information by comparing the text search on journal [10].

This research compares the two models using the Maximal Marginal Relevance (MMR) Model, where this model can summarize documents with a text calculation value seen from the level of similarity between sentences contained in a journal. [11]. Then the document is segmented into several sentences, the further process of which is revealed based on the type in viewing the summary. Furthermore, the Maximum Marginal Relevance (MMR) algorithm is also the result of an extraction method to compile a summary of one or several documents and can draw conclusions by summarizing the level of similarity between parts of the text [12]. This MMR model process avoids redundancy and selects relevant information [12].

The Maximum Marginal Relevance (MRMR) method can select sentences/text then see the results by considering the relevance between sentences and questions that combine between sentences [13]. Then the MMR algorithm can see and analyze the text results at the level of newness of information that will be used as a conclusion [14] .

Furthermore, the TextRank model in this research is an algorithm based on selecting graphs which are then processed to extract information or create a text summary from journal documents which can be retrieved on storage media which can then be used as a conclusion. [15]. The TextRank Algorithm Model can also be seen with the similarity model in the PageRank model developed by text search engines which is often seen on the Google home page which can rank web pages as a result of conclusions. [16] [17]. There is a lot of research on the TextRank algorithm used in natural language processing and text processing or summarizing text from documents which can then be used as a conclusion from various documents. [15] [18].

This research aims to compare two automatic text summarization methods, namely Maximal Marginal Relevance (MMR) and TextRank. With this comparison, research can see and evaluate the performance of both in producing text summaries that are relevant and informative [19]. Furthermore, the focus of this research is to display information from the journal conclusion process in a more organized and easy to understand way by displaying the journal title, the journal upload process which is changed to the journal text, then the results are displayed in a comparison of the two models.

Formulation of the problem

Based on the problem of this research, the following is a comparative research formula for automatic text summarization methods of maximal marginal relevance (mmr) and textrank in journals as follows :

1. How to design and implement an automatic text summarization application for journals using the Maximal Marginal Relevance (MMR) and TextRank methods?
2. How to compare the application of the Maximal Marginal Relevance (MMR) method and TextRank in automatic text summarization in journals?
3. How to see the results of the comparison between TextRank and MMR in the process of summarizing text in journals?

Materials & Methods

Research Steps

1. Data collection

   This research stage uses data obtained from journals and the journal data is entered into documents directly for automatic comparison

2. Analysis of System Requirements

   This system uses the MMR algorithm to summarize text and the TextRank algorithm carries out automatic text summarization, by comparing the Maximal Marginal Relevance algorithm with TextRank

3. System design and implementation

   The system design uses data flow diagrams and an interface to input or enter journal files in format (docx, pdf). Furthermore, the application system implementation uses the title input indicator interface, alpha (MMR), Top N, upload PDF files, journal text and the comparison results of the two models.

4. Application Testing

   Conducting a testing application by comparing the results of both Maximal Marginal Relevance (MMR) and TextRank models. the comparison results are in the form of Rank, summary, similarity, score. The results of this test are to see which value the two models are the best in seeing the automatic test summarizer.

Research methods

The research method for comparing the automatic text summarization method of maximal marginal relevance (MMR) and textrank in journals is as follows :
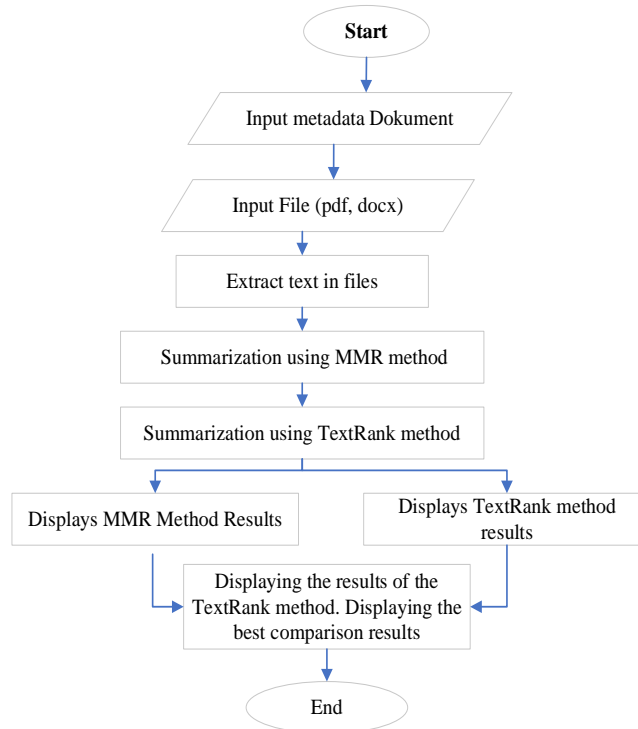
Figure 1. System Scheme

1. Input Metadata Dokumen

Input document metadata data, such as title, Alpha (MMR), Top AND, Upload pdf file, journal text into a comparison of the automatic text summarization methods of maximal marginal relevance (mmr) and textrank

2. Input File (PDF, DOCX)

   Upload PDF and Docx document files. This file can be a document in PDF, DOCX, or other text format..

3. Extraction of Text in Files

   After the file is uploaded, the application will extract text from the document so it can be extracted in a text journal

4. Summarization using the MMR Method

   MMR summarization of the extracted text. The MMR method will extract the most relevant and varied sentences to form a summary [13] .

The research method for comparing the automatic text summarization method of maximal marginal relevance (mmr) and textrank in journals is as follows:

The first step is to determine the TF-IDF method, which is a term (word) weighting method that is widely used as a comparison method against new weighting methods. [20]. In this method, the term weight calculation in a document is calculated

$$tf = 1 + Log(tf) \tag{1}$$

$$tf = 0.5 + 0.5 \, x \left(\frac{tf}{\max tf} \right) \tag{2}$$

Inverse Document Frequency (idf) calculated using a formula:

$$idf_j = \log(D/df_j) \tag{3}$$

Thus the formula for TF-IDF is a combination of the raw TF calculation formula with the IDF formula by multiplying the term frequency (tf) value by the inverse document frequency value (idf):

$$w_{ij} = tf_{ij} \; x \log (D/df_j) \tag{4}$$

Cosine Similarity

Cosine similarity is a method for calculating the distance between vectors d1 and d2 which produces the cosinex angle between the two vectors. The cosine similarity calculation equation is as follows:

$$Sim(A, B) = cosine \, \theta = \frac{A \bullet B}{||A|| \, ||B||} \tag{5}$$

Persamaan nilai AB merupakan dot product yang merupakan nilai untuk mengekspresikan sudut antara dua vector, berikut ini rumus metode mmr adalah sebagai Berikut :

$$MMR = argmax \, [\, \measuredangle * Sim1 \, S_i , Q) – (1 - \int_0) * maxSim2(S_i , S^2) \tag{6}$$

5. Summarization using the TextRank Method

   The TextRank method will be used to perform text summarization. The TextRank algorithm performs ranking calculations for each sentence in the text as follows [17]:

$$Similarity \, (S_i , S_j) \, \frac{[W_k] W_k \in S_1 \, \& \, W_k \in S_j}{log([S_1]) + log([S_1])} \tag{7}$$

Wk = There are the same number of words (terms) between sentences Si and Sj.

Si = The length of the sentence Si.

Sj = Length of sentence Sj.
The formula for PageRank is as following:

$$S(V_i) = (1 - d) + d * \sum j \in sn\ (vi) \frac{1}{[Out(Vj)]} S(Vi) \qquad (8)$$

Description:
Vi = the vertex for which the score is calculated.
Vj = vertex adjacent to Vi.
d = factors whose values are between 0 and 1.
6. Displays Summarization Results with MMR and TextRank:
   The summarization results of the two methods consist of the MMR and TextRank methods which will display the results of rank, summary, similarity and score values in the application of the automatic text summarization method maximal marginal relevance (mmr) and textrank.

Results and Discussion
A. System Analysis Maximal Marginal Relevance (MMR) and TextRank
The Maximal Marginal Relevance (MMR) system test display in looking at the comparison of the automatic text summarization method of maximal marginal relevance (mmr) and textrank in journals is as in the following table:

Table 1.Testing Maximal Marginal Relevance (MMR)

No.	Title	Alpha (MMR	Top N	Maximal Marginal Relevance (MMR)			
				Rank	Summary	Similarity	Score
1.	Pengujian Algoritma TextRank Dalam Merangkum Teks	0.7	2	1	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 man chester united bergejolak menyusul hasil minor	0.526	0.526
		0.7	2	2	Oleh karena itu untuk mengenali informasi penting dari keseluruhan isi dokumen teks dan membuat ringkasan singkat dari dokumen yang teridentifikasi Pengujian Algoritma TextRank Dalam Merangkum Teks seminar nasional Man chester United menuai rekor buruk karena kalah berturut - turut di kandang dengan skor 0-3 atau lebih untuk kali pertama dalam 61 tahun.	0.427	0.239

For TextRank testing in viewing summaries, similarity and scores in a comparison of the automatic text summarization methods of maximal marginal relevance (MMR) and TextRank in journals as in the following table:

Tabel 2. Testing TextRank

No.	Title	Alpha (MMR	Top N	Maximal Marginal Relevance (MMR)			
				Rank	Summary	Similarity	Score
1.	Pengujian Algoritma TextRank Dalam Merangkum Teks	0.7	2	1	Case Folding No Hasil Tokenizing S1 manchester united harus mengeluarkan biaya tinggi jika ingin memberhentikan sang pelatih erik ten hag S2	0.526	0.011
		0.7	2	2	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 Man chester menyusul hasil minor	0.522	0.011

Furthermore, for testing Maximal Marginal Relevance (MMR) in different titles and alpha (MMR) 0.8 and Top N totaling 5 can be seen as in the following table :

Table 3.Testing Maximal Marginal Relevance (MMR)

No.	Title	Alpha (MMR	Top N	Maximal Marginal Relevance (MMR)				
				Rank	Summary	Similarity	Score	
1.	Penerapan Algoritma Textrank Untuk Automatic Summarization Pada Dokumen Berbahasa Indonesia	0.8	5	1	Tampilan dari halaman output ini dapat dilihat pada Gambar 7 .	0.361	0.361	
		0.8	5	2	Dalam penelitian ini, akan dihasilkan ringkasan yang masih mengandung setengah kalimat yang ada dalam dokumen dan tiga perempat kalimat dari dokumen.	0.341	0.261	
		0.8	5	3	TextRank yang digunakan pada sistem ini adalah metode yang menghasilkan ekstraksi berupa kalimat (TextRank for Sentence Extraction).	0.336	0.244	
		0.8	5	4	Gambar 1 Skema sistem peringkas otomatis dengan Algoritma TextRank Kalimat yang berada di daftar urutan teratas akan dipilih menjadi ringkasan ekstraktif untuk dokumen tersebut.	0.333	0.235	
		0.8	5	5	Gambar 2 Graf yang telah dibangun dari algoritma TextRank Jurnal Ilmu Teknik dan K omputer Vol.	0.319	0.228	

For TextRank testing in viewing summaries, similarity and scores in a comparison of the automatic text summarization methods of maximal marginal relevance (MMR) and TextRank in journals as in the following table:

Tabel 4. Testing TextRank

No.	Title	Alpha (MMR	Top N	Maximal Marginal Relevance (MMR)				
				Rank	Summary	Similarity	Score	
1.	Penerapan Algoritma Textrank Untuk Automatic Summarization Pada Dokumen Berbahasa Indonesia	0.8	5	1	Tampilan dari halaman output ini dapat dilihat pada Gambar 7 .	0.361	0.007	
		0.8	5	2	TextRank yang digunakan pada sistem ini adalah metode yang menghasilkan ekstraksi berupa kalimat (TextRank for Sentence Extraction).	0.336	0.007	
		0.8	5	3	Tampilan dari halaman process ini dapat dilihat pada Gambar 6 .	0.356	0.007	
		0.8	5	4	Tampilan dari halaman help ini dapat dilihat pada Gambar 8 .	0.352	0.007	
		0.8	5	5	Dalam penelitian ini, akan dihasilkan ringkasan yang masih mengandung setengah kalimat yang ada dalam dokumen dan tiga perempat kalimat dari dokumen.	0.342	0.007	

B. System implementation Maximal Marginal Relevance (MMR) and TextRank

The system display in viewing the comparison of the automatic text summarization methods of maximal marginal relevance (mmr) and textrank in journals is as follows :

A. list of documents

The following is a list of comparative documents for automatic text summarization methods of maximal marginal relevance (mmr) and textrank in journals, as follows:
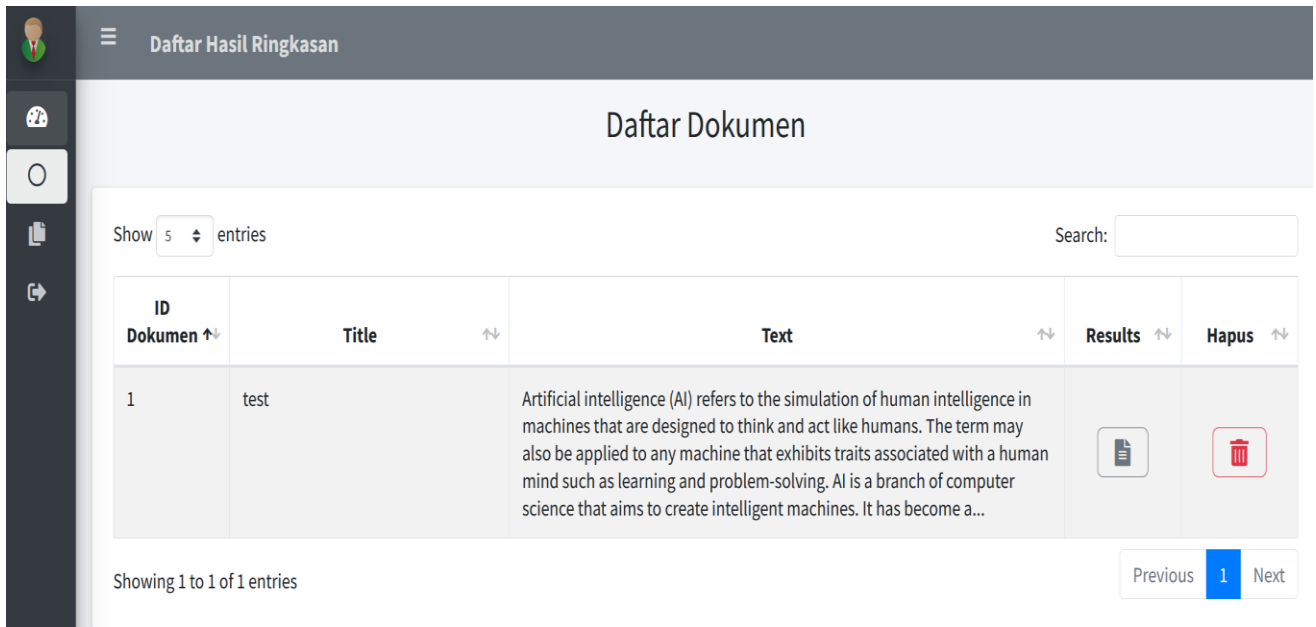
Figure 1. list of documents

B. Process MMR dan TextRank

The following is a comparison process for the automatic text summarization method of maximal marginal relevance (mmr) and textrank in journals as follows :
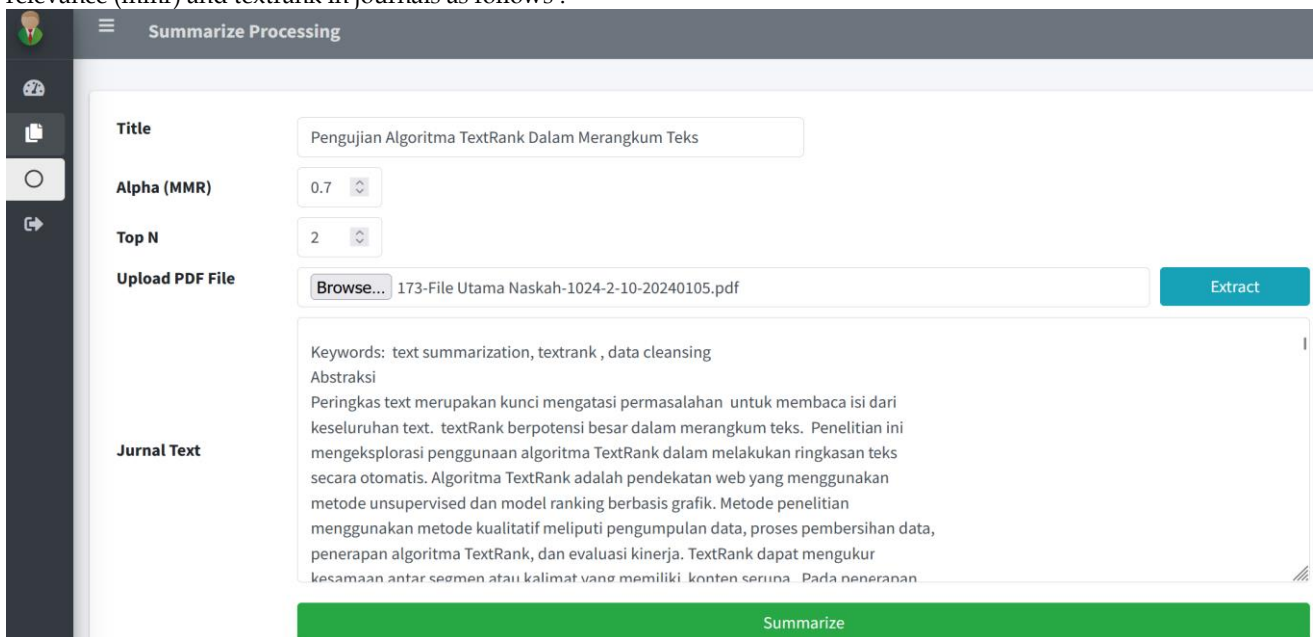

Figure 2. Process MMR dan TextRank

C. Implementasion System method of maximal marginal relevance (mmr) and textrank in journals as follows :

The following are the results of applying the comparison of the automatic text summarization method of maximal marginal relevance (mmr) and textrank in journals in the following image:

Upload PDF File | Browse... 173-File Utama Naskah-1024-2-10-20240105.pdf | Extract

Jurnal Text

Keywords: text summarization, textrank , data cleansing
Abstraksi
Peringkas 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

Summarize

Maximal Marginal Relevance (MMR)				TextRank			
Rank	Ringkasan	Similarity	Score	Rank	Ringkasan	Similarity	Score
1	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 man chester united bergejolak menyusul hasil minor yang diperoleh setan merah saat melawan newcastle dalam 16 besar piala liga inggris 2023 - 2024 S3 man chester hag di Pengujian Algoritma TextRank Dalam Merangkum Teks SEMINAR NASIONAL AMIKOM SURAKARTA (SEMNASA) 2023 e-ISSN: 3031 -5581 Sukoharjo, 25 November 2023 32 No Hasil Tokenizing united besutan erik ten hag menelan kekalahan 0-3 dari newcastle stadion old trafford S4 ini menjadi kekalahan kedua secara beruntun man chester united di stadion old trafford S5 sebelumnya, setan merah dibekuk man city 0-3 dalam lanjutan premier league S6 menurut squawka, man chester united menuai rekor buruk karena kalah berturut - turut di kandang dengan skor 0-3 atau lebih untuk kali pertama dalam 61 tahun.	0.526	0.526	1	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 man chester united bergejolak menyusul hasil minor yang diperoleh setan merah saat melawan newcastle dalam 16 besar piala liga inggris 2023 - 2024 S3 man chester hag di Pengujian Algoritma TextRank Dalam Merangkum Teks SEMINAR NASIONAL AMIKOM SURAKARTA (SEMNASA) 2023 e-ISSN: 3031 -5581 Sukoharjo, 25 November 2023 32 No Hasil Tokenizing united besutan erik ten hag menelan kekalahan 0-3 dari newcastle stadion old trafford S4 ini menjadi kekalahan kedua secara beruntun man chester united di stadion old trafford S5 sebelumnya, setan merah dibekuk man city 0-3 dalam lanjutan premier league S6 menurut squawka, man chester united menuai rekor buruk karena kalah berturut - turut di kandang dengan skor 0-3 atau lebih untuk kali pertama dalam 61 tahun.	0.526	0.011
2	Oleh karena itu untuk mengenali informasi penting dari keseluruhan isi dokumen teks dan membuat ringkasan singkat dari dokumen yang teridentifikasi Pengujian Algoritma TextRank Dalam Merangkum Teks SEMINAR NASIONAL AMIKOM SURAKARTA (SEMNASA) 2023 e-ISSN :3031 -5581 Sukoharjo, 25 November 2023 26 tersebut diperlukan sistem yang dapat secara otomatis mengolah inti dari informasi dokumen teks [2].	0.427	0.239	2	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 Man chester menyusul hasil minor yang diperoleh Setan Merah Saat Newcastle dalam 16 besar Piala Liga Inggris 2023 -2024 Pengujian Algoritma TextRank Dalam Merangkum Teks SEMINAR NASIONAL AMIKOM SURAKARTA (SEMNASA) 2023 e-ISSN :3031 -5581 Sukoharjo, 25 November 2023 31 No Hasil Tokenizing United bergejolak melawan S3 Man chester United besutan Erik ten Hag menelan kekalahan 0-3 Dari Newcastle Di Stadion Old Trafford S4 Ini menjadi kekalahan kedua secara beruntun Man chester United di Stadion Old Trafford S5 Sebelumnya, Setan Merah dibekuk Man chester City 0-3 dalam lanjutan Premier League S6 Menurut Squawka, Man chester United menuai rekor buruk karena kalah berturut - turut di kandang dengan skor 0-3 atau lebih untuk kali pertama dalam 61 tahun.	0.522	0.011
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 11 1 kata mampu diringkas hingga 74 kata.	0.387	0.234	3	Oleh karena itu untuk mengenali informasi penting dari keseluruhan isi dokumen teks dan membuat ringkasan singkat dari dokumen yang teridentifikasi Pengujian Algoritma TextRank Dalam Merangkum Teks SEMINAR NASIONAL AMIKOM SURAKARTA (SEMNASA) 2023 e-ISSN :3031 -5581 Sukoharjo, 25 November 2023 26 tersebut diperlukan sistem yang dapat secara otomatis mengolah inti dari informasi dokumen teks [2].	0.427	0.009

Simpan

Conclusions

Based on the results and discussion of the research, the following conclusions can be drawn from the Comparison Of Automatic Maximal Marginal Relevance (Mmr) And Textrank Text Summarizing Methods In Journals as follows:

1. Based on the comparison results between the MMR (Maximal Marginal Relevance) and TextRank methods, MMR is superior in seeing specific and unique information in a summary. furthermore MMR avoids redundancy and adds new information or perspectives to the summary. This is important for specific details or facts, while TextRank is more precise in providing more general information or broad context. The next comparison result MMR can match the closest relationship with the text state relationship, but there is a high possibility of losing the text picture when focusing on the summary details in a text, then textrank captures the important points that are repeated, ensuring that key aspects of the text are represented, but can be redundant.

2. In conclusion, the choice between MMR and TextRank depends on the specific goals of summarization and the type of text in a problem, the method with MMR is more appropriate for detail and specificity, while TextRank is better for general description and context.

3. The test results in the title display Testing the TextRank Algorithm in Summarizing Text, Alpha (MMR) 0.6 and Top N (0.3). summarize on the Maximal Marginal Relevance (MMR) method rank 1 for Case Folding No Results Tokenizing S1 Manchester United has to pay high costs if they want to dismiss coach Erik Ten Hag S2 Erik Ten Hag's leadership at Manchester United is in turmoil following minor results and similarity values of 0.526 score 0.526. The second rank of the text rank method results for the summarization process from the initial number of words of 11 1 words can be summarized to 74 words. the similarity results are 0.526 and the score is 0.526.

Acknowledgments

References

[1]   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)*, Aug. 2019, pp. 1–6. doi: 10.1109/IC3.2019.8844939.

[2]   R. Robiyanto, N. Nugraha, and I. Apriatna, "Peringkasan Teks Otomatis Berita Menggunakan Metode Maximum Marginal Relevance," *JEJARING J. Teknol. dan Manaj. Inform.*, vol. 4, no. 1, pp. 23–32, May 2019, doi: 10.25134/jejaring.v4i1.6712.

[3]   Y. K. Khor, C. W. Tan, and T. M. Lim, "Text Summarization on Amazon Food Reviews using TextRank," in *Conference Proceedings: International Conference on Digital Transformation and Applications (ICDXA 2021)*, 2021, pp. 113–120. doi: 10.56453/icdxa.2021.1011.

[4]   R. T. Adek, Bustami, and M. Ula, "Systematics Review on the Application of Social Media Analytics for Detecting Radical and Extremist Group," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1071, no. 1, p. 012029, Feb. 2021, doi: 10.1088/1757-899X/1071/1/012029.

[5]   N. Nurdin, R. Rizal, and R. Rizwan, "Pendeteksian Dokumen Plagiarisme dengan Menggunakan Metode Weight Tree," *Telematika*, vol. 12, no. 1, p. 31, Feb. 2019, doi: 10.35671/telematika.v12i1.775.

[6]   M. F. Hidayattullah and A. Azizi, "Peringkas 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.

[7]   A. M. A. Zeyad and A. Biradar, "Abstractive Multi-Document Summarization: Exploiting Maximal Marginal Relevance and Pretrained Models," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Jul. 2023, pp. 1–5. doi: 10.1109/ICCCNT56998.2023.10307351.

[8]   I. M. Karo Karo, S. Dewi, and A. Perdana, "Implementasi Text Summarization Pada Review Aplikasi Digital Library System Menggunakan Metode Maximum Marginal Relevance," *JEKIN - J. Tek. Inform.*, vol. 4, no. 1, pp. 25–31, Apr. 2024, doi: 10.58794/jekin.v4i1.671.

[9]   D. P. Purbawa, Malikhah, R. N. Esti Anggraini, and R. Sarno, "Automatic Text Summarization using Maximum Marginal Relevance for Health Ethics Protocol Document in Bahasa," in *2021 13th International Conference on Information & Communication Technology and System (ICTS)*, Oct. 2021, pp. 324–329. doi: 10.1109/ICTS52701.2021.9607951.

[10]   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.

[11]   S. I. K and S. R. Balasundaram, "Phrase Embedding Based Multi Document Summarization with Reduced Redundancy using Maximal Marginal Relevance," in *2020 International Conference on Electrical Engineering and Informatics (ICELTICs)*, Oct. 2020, pp. 1–5. doi: 10.1109/ICELTICs50595.2020.9315474.

[12]   D. F. AL-Hafiidh, I. F. Rozi, and I. K. Putri, "Peringkasan Teks Otomatis pada Portal Berita Olahraga menggunakan metode Maximum Marginal Relevance.," *J. Inform. Polinema*, vol. 8, no. 3, pp. 21–30, Jun. 2022, doi: 10.33795/jip.v8i3.519.

[13]   R. Kurmi and P. Jain, "Text summarization using enhanced MMR technique," in *2014 International Conference on Computer Communication and Informatics*, Jan. 2014, pp. 1–5. doi: 10.1109/ICCCI.2014.6921769.

[14]   Shasha Xie and Yang Liu, "Using corpus and knowledge-based similarity measure in Maximum Marginal Relevance for meeting summarization," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, Mar. 2008, pp. 4985–4988. doi: 10.1109/ICASSP.2008.4518777.

[15]   S. Yu, J. Su, P. Li, and H. Wang, "Towards High Performance Text Mining," *Int. J. Grid High Perform. Comput.*, vol. 8, no. 2, pp. 58–75, Apr. 2016, doi: 10.4018/IJGHPC.2016040104.

[16]   A. Babu G. L. and S. Badugu, "Extractive Summarization of Telugu Text Using Modified Text Rank and Maximum Marginal Relevance," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 9, pp. 1–18, Sep. 2023, doi: 10.1145/3600224.

[17]   S. Zaware, D. Patadiya, A. Gaikwad, S. Gulhane, and A. Thakare, "Text Summarization using TF-IDF and Textrank algorithm," in *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*, Jun. 2021, pp. 1399–1407. doi: 10.1109/ICOEI51242.2021.9453071.

[18]   U. Rani and K. Bidhan, "Comparative Assessment of Extractive Summarization: TextRank, TF-IDF and LDA," *J. Sci. Res.*, vol. 65, no. 01, pp. 304–311, 2021, doi: 10.37398/JSR.2021.650140.

[19]   G. Sharma and D. Sharma, "Automatic Text Summarization Methods: A Comprehensive Review," *SN Comput. Sci.*, vol. 4, no. 1, p. 33, Oct. 2022, doi: 10.1007/s42979-022-01446-w.

[20]   R. T. Adek and M. Ula, "A Survey on The Accuracy of Machine Learning Techniques for Intrusion and Anomaly Detection on Public Data Sets," in *2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA)*, Jul. 2020, pp. 19–27. doi: 10.1109/DATABIA50434.2020.9190436.