Evaluasi Kinerja TextRank dan LexRank Berbasis TF-IDF dan Word2vec untuk Text Summarization

Albertin Caecilia Djema[a1], I Gusti Ngurah Anom Cahyadi Putra[a2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana, Bali
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
[1]djema.2308561056@student.unud.ac.id
[2]anom.cp@unud.ac.id

Abstract

Text summarization is a crucial task in natural language processing, aiming to extract essential information from lengthy texts. The choice of summarization method significantly influences the quality of the generated summary. This study evaluates the performance of the TextRank and LexRank algorithms, both combined with TF-IDF and Word2Vec-based word representation techniques. The IndoSum dataset was used as the benchmark, with preprocessing steps including text cleaning, case folding, tokenization, and vector transformation using Word2Vec and TF-IDF weighting. The ROUGE metric was employed to assess summarization quality. Experimental results indicate that the TextRank algorithm, when integrated with TF-IDF and Word2Vec, achieves higher ROUGE scores compared to LexRank in generating extractive summaries.

Keywords: *TextRank, LexRank, Text Summarization, TF-IDF, Word2vec, Extractive Summary.*

1. Pendahuluan

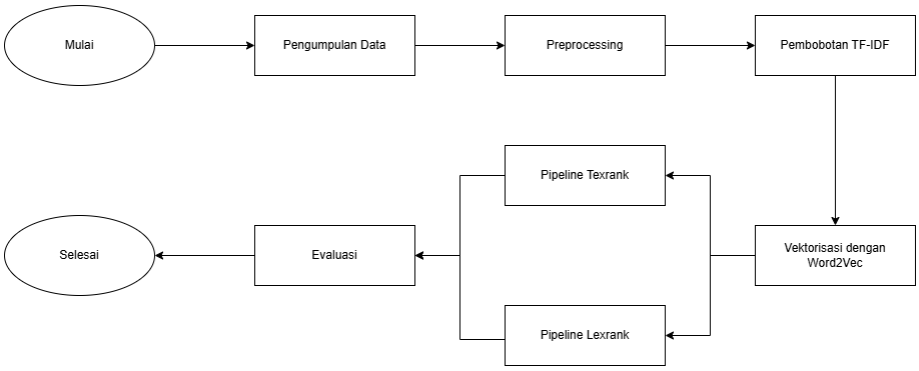
Berdasarkan laporan yang dituliskan oleh UNESCO, dikatakan bahwa minat baca masyarakat Indonesia sangatlah rendah yaitu berada di angka 0,001% yang berarti bahwa dari 1000 orang masyarakat Indonesia, hanya ada 1 orang yang rajin membaca. Namun, melihat perubahan perilaku masyarakat di era digital minat membaca tidak melulu diukur dari seberapa banyak buku yang dibaca tetapi juga seberapa banyak sumber informasi atau bahan bacaan daring yang dibaca, dibagikan, didiskusikan, disimpan atau diunduh [2]. Dengan ini media sosial yang memiliki karakteristik yang unik dapat dimanfaatkan untuk meningkatkan minat baca masyarakat Indonesia yang sangat rendah. Hal ini bukanlah tidak mungkin mengingat pada Januari 2025 tercatat bahwa 50,2% dari jumlah total populasi masyarakat Indonesia yaitu sebanyak 143 juta identitas aktif menggunakan media sosial [3]. Di satu sisi hal ini bisa dimanfaatkan untuk menghidupkan kembali kebiasaan membaca dengan memanfaatkan kecenderungan masyarakat untuk mengaktualisasi diri di media sosial, di sisi lain pengguna media sosial cenderung mengambil informasi secara instan dan tidak menginginkan proses panjang seperti membaca berita secara keseluruhan dan memeriksa keakuratan informasi[2]. Masyarakat cenderung menghindari membaca teks panjang secara keseluruhan namun tetap ingin mengambil informasi dari bacaan panjang tersebut. Kecenderungan ini kemudian bisa mengarah kepada misinformasi.

Tantangan utama yang muncul adalah bagaimana informasi atau bacaan dengan teks yang panjang tersebut dapat diringkas atau diambil inti sarinya sehingga para pembaca hanya perlu membaca teks yang lebih singkat tapi tidak menghilangkan informasi penting sehingga tidak berujung pada misinformasi. *Text Summarization* merupakan salah satu solusi yang dikembangkan dalam *Natural Language Processing* (NLP) yaitu sebuah proses otomatis untuk menghasilkan ringkasan dari sebuah sumber berupa teks. Algoritma yang digunakan dalam text summarization terdiri atas berbagai algoritma. Dengan banyaknya algoritma dari text summarization ini maka penelitian ini akan mengevaluasi kinerja kerja dari dua algoritma text summarization yaitu TextRank dan LexRank dengan mengkombinasikannya dengan TF-IDF dan Word2Vec dalam merepresentasi data teks. Tujuan penelitian ini adalah untuk mengetahui dan

membandingkan kinerja kedua algoritma tersebut apabila dikombinasikan dengan TF-IDF dan Word2Vec. Metode evaluasi yang digunakan adalah Rouge-N. Adapun penelitian serupa yang sudah dilakukan sebelumnya yaitu pada tahun 2022 yaitu sebuah penelitian yang dilakukan oleh Satya Deo dan Debajyoty Banik yang berjudul Text summarization using Texrank and LexRank through Latent Semantic Analysis [4]. Penelitian ini bertujuan membandingkan performa metode TextRank, LexRank, dan Latent Semantic Analysis (LSA) dalam merangkum teks dari berbagai domain blog. Evaluasi dilakukan menggunakan metrik cosine similarity, unit overlap, dan recall. Hasilnya menunjukkan bahwa TextRank menghasilkan ringkasan paling baik secara keseluruhan, sementara LSA unggul dalam metrik cosine similarity dan LexRank menunjukkan performa paling rendah [4]. Penelitian lainnya juga dilakukan pada tahun 2023 dimana penelitian ini mengusulkan pengembangan algoritma TextRank dengan menggabungkan teknik weighted word embedding untuk meningkatkan kualitas *text summarization* [5]. Dalam penelitian ini, penulis menggunakan tiga model representasi kata: Word2Vec, FastText, dan IndoBERT, serta mengintegrasikan bobot TF-IDF untuk membentuk representasi kalimat yang lebih akurat. Hasil eksperimen menunjukkan bahwa penggunaan embedding berbobot secara signifikan meningkatkan kinerja TextRank dibandingkan metode dasarnya, dengan peningkatan skor Rouge-1 hingga 17,33% dan Rouge-2 hingga 30,01%. Metode ini terbukti paling efektif saat menggunakan embedding dari model kontekstual seperti IndoBERT, namun TF-IDF memberikan dampak paling besar pada model berbasis embedding statis seperti Word2Vec dan FastText. Berdasarkan kajian tersebut terdapat perbedaan pada penelitian yang akan yang akan dilakukan oleh penulis yaitu pada penelitian ini penulis akan membandingkan algoritma TextRank dan Lexrank yang dikombinasikan dengan TF-IDF dan Word2Vec dan kemudian dievaluasi menggunakan Rouge-N.

2. Metode Penelitian

Metode penelitian yang akan dilakukan terdiri atas beberapa tahap yang terurut sehingga mampu mendapatkan hasil yang maksimal serta memenuhi tujuan penelitian yang sudah ditetapkan. Berikut merupakan gambaran alur penelitian yang akan dilakukan.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diambil dari IndoSum, yang merupakan sebuah benchmark dataset summarization berbahasa indonesia yang bersifat open acces[6]. Dataset ini diunduh melalui tautan google drive yang dibagikan secara umum oleh pemiliknya. File dalam dataset ini memiliki ekstensi *.jsonl*, yang berisi artikel berita lengkap dengan ringkasan ekstraktifnya sehingga bisa digunakan sebagai ground truth dalam proses evaluasi. Jumlah data yang digunakan yaitu sebanyak 71.353 artikel. Struktur data dapat dilihat pada tabel 1.

Tabel 1. Key Data

Key	Keterangan Key
Id	Pengenal unik untuk membedakan satu artikel dengan artikel lainnya
Paragraphs	Isi artikel
Summary	Ringkasan abstraktif, yang terdiri atas list kalimat
Gold Labels	Label kalimat yang menjadi bagian dari ringkasan ekstraktif
category	Kategori artikel

2.2. Preprocessing

Preprocessing merupakan sebuah tahapan yang dilakukan untuk menyiapkan data sebelum dimasukan ke model utama sehingga data dapat lebih mudah untuk diolah. Pada penelitian ini dilakukan beberapa tahapan preprocessing dimulai dari case folding sampai pada stemming

a. Cleaning Data

Pada penelitian ini tahap pertama yang dilakukan dalam preprocessing adalah cleaning data yaitu menghapus karakter-karakter tidak penting yang terdapat dalam artikel sehingga tidak mengganggu proses representasi

b. Case Folding

Case folding adalah proses menyeragamkan semua huruf yang ada dalam kalimat dengan mengubahnya menjadi huruf kecil atau lowercase. Hal ini dilakukan agar kedua kata yang sebenarnya sama tidak dinilai sebagai dua kata yang berbeda hanya karena terdapat perbedaan penempatan huruf kapital antara kedua kata tersebut.

c. Tokenisasi

Proses berikutnya, yang sekaligus menjadi proses terakhir adalah tokenisasi yaitu sebuah proses untuk memecah artikel menjadi potongan-potongan kata. Proses ini bertujuan mempermudah perhitungan bobot setiap kata untuk kemudian membangun kalimat dan pada akhirnya artikel

2.3 Pembobotan TF-IDF

TF-IDF (Term Frequency-Inverse Document) merupakan salah satu metode pembobotan yang sering digunakan dalam berbagai penelitian, teknik pembobotan ini sering digunakan karena implementasinya yang terbilang mudah namun juga cukup kompleks untuk merepresentasikan bobot dari data sehingga mampu meningkatkan keakuratan suatu algoritma[7]. TF-IDF pada dasarnya merupakan suatu metode pemberian bobot terhadap hubungan antara kata dengan suatu dokumen dimana metode ini terdiri atas dua konsep perhitungan yaitu menggabungkan frekuensi kemunculan kata dalam dokumen tertentu dengan frekuensi kebalikan dari dokumen yang mengandung kata tersebut[sintia sarjon]. Pentingnya suatu kata dalam dokumen tertentu dapat dilihat dari frekuensi kemunculan kata tersebut dalam dokumen. Adapun rumus perhitungan yang dilakukan oleh TF-IDF dapat dilihat dalam persamaan 1[8].

$$W(d,t) = TF(d,t) \times log\ log\ \left(\frac{N}{df(t)}\right) \quad (1)$$

$$TF(d,t) = \frac{jumlah\ kata\ dalam\ dokumen}{total\ kata\ dalam\ dokumen} \quad (2)$$

2.4 Word2vec

Setelah melakukan pembobotan dengan tf-idf, tahap berikutnya adalah membentuk vector kata dengan menggunakan word2vec sebagai representasi semantik dari kata-kata dalam artikel. Model yang digunakan adalah FastText versi Bahasa Indonesia (cc.id.300.v3c). Setiap kata direpresentasikan oleh word2vec sebagai vektor berdimensi tetap (300 dimensi) Vektor kata ini yang kemudian akan digunakan untuk membentuk representasi kalimat dan juga pada akhirnya artikel dengan cara menghitung rata-rata dari vektor kata yang ada dengan tetap mempertimbangkan bobot penting kata berdasarkan TF-IDF. Adapun persamaan yang digunakan untuk membuat vektor kalimat $\underline{s}$ dapat dilihat dalam persamaan (3).

$$\underline{s} = \frac{1}{N}\sum_{i=1}^{N} \quad w_i \cdot \underline{v_i} \tag{3}$$

Keterangan:

$N = jumlah\ kata\ dalam\ kalimat$
$w_i = vektor\ kata\ ke-1\ daro\ model$
$\underline{v_i} = bobot\ TF-IDF\ dari\ kata\ ke-i$

2.5 TextRank

Algoritma TextRank merupakan salah satu algoritma peringkas teks berbasis graf yang termasuk dalam kategori unsupervised learning, dimana algoritma ini merangkum teks dengan memilih kalimat-kalimat yang memiliki skor tertinggi untuk kemudian dimasukkan ke dalam ringkasan akhir[4]. Dalam penelitian ini dikarenakan ringkasan yang dihasilkan merupakan ringkasan ekstraktif maka algoritma textrank akan membentuk graf setiap simpul mewakili kalimat, dan sisi-sisi antara simpul menggambarkan tingkat kemiripan antar kalimat tersebut. Kemiripan kata dalam penelitian ini dihitung menggunakan cosine similarity.

2.6 LexRank

Sebagai perbandingan algoritma lain yang digunakan dalam penelitian ini adalah Algoritma LexRank. Algoritma LexRank pada umumnya memiliki prinsip yang sama dengan TextRank yaitu sebuah algoritma peringkas text berbasis graf dengan simpul sebagai representasi kalimat dan sisi atau edge yang mewakili kemiripan antara kalimat dengan menghitung cosine similarity antar setiap kalimat yang ada[4]. Perbedaan antara kedua algoritma ini adalah bahwa Lexrank memiliki fitur untuk mengatur threshold dari kemiripan antara setiap kalimat sehingga graf yang terbentuk cenderung lebih jarang daripada graf yang dibentuk oleh TextRank dengan data yang sama.

2.7 Evaluasi

Metrik evaluasi yang digunakan dalam penelitian ini adalah ROUGE atau recall-oriented understudy for gisting evaluation. ROUGE merupakan salah satu metrik yang sering digunakan untuk menilai hasil ringkasan ekstraktif dalam pengembangan sistem peringkas teks otomatis. Secara umum metrik ini mengukur seberapa mirip hasil ringkasan yang dihasilkan oleh sistem dengan ringkasan yang dianggap benar pada dataset (ground truth) dimana pengukuran ini dilakukan dengan cara melihat kesamaan kata di antara kedua ringkasan[5]. Adapun jenis ROUGE yang digunakan dalam penelitian ini adalah ROUGE-N dan ROUGE-L. ROUGE-N digunakan untuk menghitung jumlah potongan kata berurutan (n-gram) yang sama antara kedua ringkasan sedangkan ROUGE-L menilai urutan kata terpanjang yang muncul dalam kedua ringkasan. Dalam penelitian ini F1 digunakan sebagai penilaian akhir, namun nilai Precision dan Recall juga akan ditampilkan. Adapun cara menghitung F1, Precision, dan Recall dari ROUGE-N dan ROUGE-L dapat dilihat pada persamaan (4) sampai (9).

$$Precision_{ROUGE-N} = \frac{Jumlah\ n-gram\ yang\ cocok}{Jumlah\ total\ n-gram\ dalam\ ringkasan\ sistem} \tag{4}$$

$$Recall_{ROUGE-N} = \frac{Jumlah\ n-gram\ yang\ cocok}{Jumlah\ total\ n-gram\ dalam\ ringkasan\ referensi} \quad (5)$$

$$F1_{ROUGE-N} = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \quad (6)$$

$$Precision_{ROUGE-L} = \frac{Panjang\ LCS}{Jumlah\ kata\ dalam\ ringkasan\ sistem} \quad (7)$$

$$Recall_{ROUGE-L} = \frac{Panjang\ LCS}{Jumlah\ kata\ dalam\ ringkasan\ referensi} \quad (8)$$

$$F1_{ROUGE-L} = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \quad (9)$$

3. Hasil dan Pembahasan

Penelitian ini bertujuan untuk mengevaluasi dua algoritma yang berbeda dalam sistem peringkasan teks berbasis graf yaitu algoritma TextRank dan LexRank. Adapun jenis ringkasan yang dihasilkan adalah ringkasan ekstraktif yaitu mengambil kalimat paling penting dan paling mewakili informasi dalam artikel. Perbedaan antara kedua algoritma ini terletak pada cara kedua algoritma membangun graf. Dalam algoritma TextRank graf yang dibangun adalah fully connected graph dimana semua kalimat saling berhubungan berapapun nilai kemiripan antara kalimat. Sedangkan dalam algoritma LexRank tidak semua simpul (kalimat) saling terhubung dikarenakan adanya threshold yang ditetapkan dalam algoritma ini. Adapun Threshold yang digunakan dalam penelitian ini adalah sebesar 0.3.

3.1 Pengumpulan data

Dataset dengan jumlah 71.353 akan digunakan seluruhnya dalam penelitian ini, kemudian dari beberapa key yang terdapat dalam dataset hanya akan ada dua key yang digunakan yaitu ‘paragraph’ dan ‘gold labels’. Hal ini dikarenakan ground truth yang akan digunakan adalah ringkasan ekstraktif dari dataset tersebut. Karen sebelumnya ‘gold labels’ hanya berupa list dari label kalimat yang termasuk dalam ringkasan ekstraktif maka akan diubah ke dalam list kalimat, sehingga mendapat data seperti pada tabel 2.

Tabel 2. Dataset

No	Paragraph	Gold Labels
1.	**Jakarta, CNN Indonesia** – Dokter Ryan Thamrin, yang terkenal lewat acara *Dokter Oz Indonesia*, meninggal dunia pada Jumat (4/8) dini hari. Dokter Lula Kamal yang merupakan selebritis sekaligus rekan kerja Ryan menyebut kawannya itu sudah sakit sejak setahun yang lalu. Lula menuturkan, sakit itu membuat Ryan mesti vakum dari semua kegiatannya, termasuk menjadi pembawa acara *Dokter Oz Indonesia*. Kondisi itu membuat Ryan harus kembali ke kampung halamannya di Pekanbaru, Riau, untuk menjalani istirahat.…….…….	dokter lula kamal yang merupakan selebriti sekaligus rekan kerja ryan menyebut kawannya itu sudah sakit sejak setahun yang lalu lula menuturkan sakit itu membuat ryan mesti vakum dari semua kegiatannya termasuk menjadi pembawa acara dokter oz indonesia kondisi itu membuat ryan harus kembali ke kampung halamannya di pekanbaru riau untuk menjalani istirahat
2.	**Suara.com** – Pengadilan Tindak Pidana Korupsi menggelar sidang lanjutan kasus korupsi proyek pengadaan e-KTP dengan terdakwa **Setya Novanto**, Senin (26/2/2018). Sidang itu beragenda pemeriksaan saksi. Jaksa Penuntut Umum menghadirkan tujuh orang	suaracom pengadilan tindak pidana korupsi menggelar sidang lanjutan kasus korupsi proyek pengadaan e ktp dengan terdakwa setya

p-ISSN: 2986-3929
e-ISSN: 3032-1948

No	Paragraph	Gold Labels
	saksi. Salah satunya adalah pengacara kondang, **Elza Syarief**. “Saudara Elza Syarief, pekerjaan advokat,” kata hakim Yanto saat menanyakan identitas saksi di Gedung Pengadilan Tipikor, Jalan Bungur Besar Raya, Kemayoran, Jakarta Pusat……..	novanto senin 26 2 2018 jaksa penuntut umum menghadirkan tujuh orang saksi salah satunya pengacara kondang elza syarief sementara enam saksi yang lainnya adalah pensiunan pns atau mantan kabag umum kemendagri rudy endarto hakim kepala biro perlengkapan sekjen kemendagri yudi pramadi dan kepala tim teknis proyek e ktp husni fahmi

3.2 Preprocessing

Data yang sudah dikumpulkan kemudian melalui preprocessing dengan tujuan untuk menyeragamkan data dan memudahkan data untuk diproses oleh model. Hasil preprocessing dapat dilihat pada tabel 3. Dengan hasil pada tabel hanya mencakup kalimat pertama dalam salah satu artikel saja.

Tabel 3. Hasil Preprocessing

No	Tahapan	Hasil
1.	Data Awal	Pengadilan tindak pidana korupsi menggelar sidang lanjutan kasus korupsi proyek pengadaan E-KTP dengan terdakwa Setya Novanto Senin, 26-2-2018
2.	Cleaning	Pengadilan tindak pidana korupsi menggelar sidang terdakwa kasus korupsi proyek pengadaan E KTP dengan terdakwa Setya Novanto Senin
3.	Case Folding	Pengadilan tindak pidana korupsi menggelar sidang terdakwa kasus korupsi proyek pengadaan e ktp dengan terdakwa setya novanto senin
4.	T Tokenisasi	[‘Pengadilan’, ‘tindak’, ‘pidana’, ‘korupsi’, ‘menggelar’, ‘sidang’, ‘terdakwa’, ‘kasus’, ‘korupsi’, ‘proyek’, ‘pengadaan’, ‘e’, ‘ktp’, ‘dengan’, ‘terdakwa’, ‘setya’, ‘novanto’, ‘senin’]

3.3 TF-IDF

Setelah melalui preprocessing hal pertama yang akan dilakukan adalah pembobotan dengan TF-IDF, dengan persamaan yang sudah dijelaskan sebelumnya. Hasil pembobotan kata dalam satu artikel dapat dilihat pada gambar 2.

	Kata	TF-IDF
0	ryan	0.514330
1	lula	0.447335
2	oz	0.293480
3	dokter	0.254927
4	penyakit	0.148426
5	enggak	0.125649
6	jenguk	0.120508
7	sakit	0.119991
8	penyebab	0.111843
9	barangkali	0.110558

Gambar 2. Bobot Kata

p-ISSN: 2986-3929
e-ISSN: 3032-1948

3.4 Word2Vec

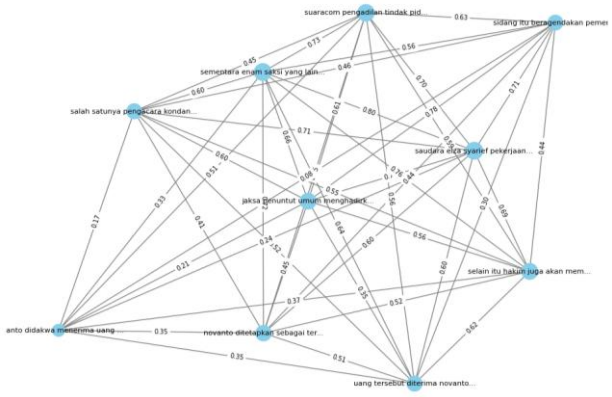
Setelah melakukan pembobotan terhadap setiap kata dan dokumen menggunakan TF-IDF sistem selanjutnya membentuk vektor kata yang kemudian membentuk kalimat dan pada akhirnya adalah artikel. Hasil perhitungan vektor kalimat dapat dilihat pada gambar 3.

	Kalimat	Dimensi_0	Dimensi_1	Dimensi_2	Dimensi_3	Dimensi_4	Dimensi_5	Dimensi_6	Dimensi_7	Dimensi_8	...
0	jakarta cnn indonesia dokter ryan thamrin yang...	0.182933	0.121991	-0.118030	0.086225	-0.058555	0.050085	0.235205	0.396617	0.027869	...
1	dokter lula kamal yang merupakan selebriti sek...	-0.009863	-0.071374	0.106710	0.309664	-0.057413	-0.303980	0.027067	-0.021608	0.007587	...
2	lula menuturkan sakit itu membuat ryan mesti v...	0.127290	0.107613	-0.039919	0.108403	-0.079258	-0.004260	0.244442	0.357203	0.040256	...
3	kondisi itu membuat ryan harus kembali ke	0.008604	-0.019295	0.065578	0.250420	-0.005637	-0.242829	-0.086133	0.029721	-0.088005	...

Gambar 3. Vektor kalimat

3.5 TextRank

Setelah melakukan pembobotan dengan TF-IDF dan membangun vektor kata menggunakan Word2vec, algoritma textrank kemudian menghitung kemiripan antara setiap kalimat kemudian membangun fully connected graph dengan menggunakan kalimat sebagai simpul serta nilai kemiripan antar kalimat sebagai bobot edge. Graf yang dibangun oleh TextRank dapat dilihat pada gambar 4.



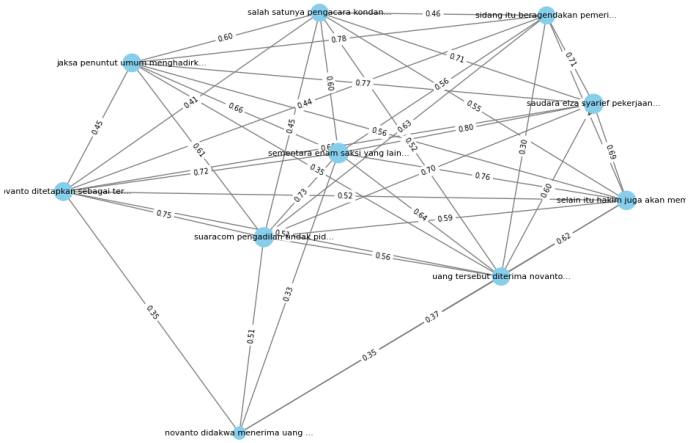
Gambar 4 . Graf TextRank

Dapat dilihat bahwa graf yang dibangun oleh TextRank sangat padat, ini dikarenakan pada algoritma TextRank semua kalimat terhubung berapapun nilai kemiripan yang mereka miliki.

3.6 LexRank

Sama seperti TextRank, algoritma LexRank juga membentuk graf namun tidak semua kalimat terhubung dalam graf ini. Hanya graf dengan nilai kemiripan di atas threshold (dalah penelitian ini 0.3) yang dapat dihubungkan dengan edge. Graf yang dibangun oleh LexRank dapat dilihat pada gambar 5.

p-ISSN: 2986-3929
e-ISSN: 3032-1948



Gambar 5 . Graf LexRank

Graf yang terbentuk tidak sepadat graf pada algoritma TextRank. Bisa dilihat bahwa tidak semua edge terhubung antara satu sama lain, serta bobot edge yang ada juga hanya edge dengan bobot lebih dari 0.3 (threshold).

3.7 Evaluasi

Ringkasan yang dihasilkan oleh sistem kemudian dibandingkan dengan ringkasan yang dianggap benar pada dataset (ground truth) lalu dilakukan evaluasi dengan metrik ROUGE-N dan ROUGE-L. Hasil evaluasi ini dapat dilihat pada tabel 3.

Tabel 5. Hasil Evaluasi TextRank

Metrik Evaluasi	Precision	Recall	F-1 Score
ROUGE-1	0.4579	0.4515	0.4444
ROUGE-2	0.3291	0.3280	0.3212
ROUGE-L	0.3570	0.3291	0.3474

Tabel 6. Hasil Evaluasi LexRank

Metrik Evaluasi	Precision	Recall	F-1 Score
ROUGE-1	0.4266	0.4412	0.4255
ROUGE-2	0.3027	0.3142	0.3025
ROUGE-L	0.3312	0.3438	0.3320

Berdasarkan hasil evaluasi yang ditampilkan pada tabel 5 dan tabel 6, diketahui bahwa algoritma TextRank secara konsisten menunjukkan performa yang lebih baik dibandingkan LexRank dalam menghasilkan ringkasan teks yang diukur menggunakan tiga metrik ROUGE: ROUGE-1, ROUGE-2, dan ROUGE-L. Pada metrik ROUGE-1, extRank memperoleh nilai precision sebesar 0.4579, recall sebesar 0.4515, dan F-1 score sebesar 0.4444. Sementara itu, LexRank mencatat precision sebesar 0.4266, recall 0.4412, dan F-1 score 0.4255. Selisih nilai F-1 score antara kedua metode pada metrik ini adalah **0.0189**, yang menunjukkan bahwa TextRank mampu mempertahankan lebih banyak informasi penting dalam ringkasan.

Pada metrik ROUGE-2, yang mengukur kesesuaian urutan pasangan kata (bigram), TextRank kembali unggul dengan F-1 score sebesar 0.3212 dibandingkan LexRank yang memperoleh nilai

p-ISSN: 2986-3929
e-ISSN: 3032-1948

sebesar 0.3025. Selisih F-1 score pada metrik ini adalah **0.0187**, mengindikasikan bahwa ringkasan hasil TextRank lebih relevan dalam konteks struktur bahasa dibandingkan LexRank. Sementara itu, pada metrik ROUGE-L, yang memperhitungkan kesamaan urutan kata terpanjang (longest common subsequence), TextRank memperoleh F-1 score sebesar 0.3474, lebih tinggi dari LexRank yang hanya mencapai 0.3320, dengan selisih sebesar **0.0154**.

Secara keseluruhan, TextRank unggul dalam semua metrik evaluasi dibandingkan LexRank, dengan selisih yang cukup konsisten pada setiap metrik, terutama pada ROUGE-1 dan ROUGE-2. Hal ini menunjukkan bahwa TextRank menghasilkan ringkasan yang lebih informatif dan lebih menyerupai ringkasan referensi dibandingkan LexRank, sehingga dapat disimpulkan bahwa dalam konteks dataset dan pengujian ini, TextRank merupakan metode yang lebih efektif.

4. Kesimpulan

Setelah melakukan penelitian dapat disimpulkan bahwa algoritma TextRank dan Lexrank yang dikombinasikan dengan representasi berbasis TF-IDF dan Word2vec mampu menghasilkan ringkasan teks ekstraktif dengan tingkat akurasi yang berbeda. Setelah melakukan evaluasi dengan metrik ROUGE, baik ROUGE-N maupun ROUGE-L terhadap data IndoSum, algoritma TextRank ternyata menghasilkan akurasi yang lebih unggul dibandingkan dengan algoritma Lexrank. Meskipun kedua algoritma ini sama-sama berbasis graf namun cara pembentukan graf yang berbeda menyebabkan hasil yang berbeda pula. Hal ini berarti bahwa struktur graf dan perhitungan bobot hubungan antar kalimat yang digunakan TextRank lebih efektif dalam mengekstraksi informasi penting.Meskipun hasil yang diperoleh cukup memuaskan, penelitian ini masih memiliki keterbatasan pada penggunaan model representasi kata yang bersifat statis. Untuk pengembangan selanjutnya, disarankan untuk mengeksplorasi penggunaan model representasi kontekstual seperti BERT, serta membandingkan dengan metode berbasis pembelajaran mendalam guna meningkatkan kualitas ringkasan dan memperluas cakupan evaluasi.

Daftar Pustaka

[1] T. Tawali, "Reading Interest of Literacy Taman Membaca Visitor at Penujak Village Lombok Tengah," *Teaching and Learning Journal of Mandalika*, vol. 5, no. 2, pp. 1–7, Nov. 2024.

[2] N. Kurniasih, “Reading Habit in Digital Era: Indonesian People do not Like Reading, is it True?”, 13-Nov-2017. [Online]. Available: osf.io/preprints/inarxiv/5apkf_v1.

[3] S. Kemp, “Digital 2025: Indonesia,” *We Are Social* and *Hootsuite*, Jan. 2025. [Online]. Available: https://andi.link/hootsuite-we-are-social-data-digital-indonesia-2025/

[4] S. Deo dan D. Banik, “Text Summarization using TextRank and LexRank through Latent Semantic Analysis,” *2022 OITS International Conference on Information Technology (OCIT)*, pp. 113–118, 2022.

[5] E. Yulianti, N. Pangestu, dan M. A. Jiwanggi, “Enhanced TextRank using weighted word embedding for text summarization,” *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 13, no. 5, pp. 5472–5482, Oct. 2023.

[6] K. Kurniawan and S. Louvan, “IndoSum: A New Benchmark Dataset for Indonesian Text Summarization,” in *Proc. 2018 Int. Conf. Asian Lang. Process. (IALP)*, Bandung, Indonesia, Nov. 2018, pp. 215–220..

[7] N. M. G. S. Nirmalaa dan N. A. S. E. R. Ngurah, “Analisis Sentimen dengan Logistic Regression untuk Deteksi Kata pada Livin’ by Mandiri,” *Jurnal Nasional Teknologi Informasi dan Aplikasinya (JNATIA)*, vol. 2, no. 4, pp. 869–875, Aug. 2024.

[8] P. W. A. Wibawa and C. R. A. Pramartha, “Analisis Sentimen pada Teks Berbahasa Bali Menggunakan Metode Multinomial Naive Bayes dengan TF-IDF dan BoW,” *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, vol. 2, no. 1, pp. 37–46, 2023.

p-ISSN: 2986-3929
e-ISSN: 3032-1948

Halaman ini sengaja dibiarkan kosong