

IMPLEMENTASI TEXT SUMMARIZATION REVIEW APLIKASI TOKOPEDIA DI GOOGLE PLAYSTORE MENGGUNAKAN ALGORITMA TEXT RANK

Bryliana Delvia Maharani¹⁾, Alexius Endi Budianto²⁾, Wahyudi Harianto³⁾

Universitas PGRI Kanjuruhan Malang^{1,2,3)}
dmbyliana@gmail.com

Abstrak

Bagaimana text summarization dalam melakukan review aplikasi tokopedia di google play store Bagaimana penerapan algoritma text rank dalam melakukan peringkasan teks pada review aplikasi Tokopedia di playstore Mengukur akurasi dan performa algoritma text rank dalam peringkasan teks pada review aplikasi tokopedia di google playstore Seberapa efektif textrank jika digunakan untuk text summarization Kerangka berpikir adalah struktur yang digunakan untuk membantu peneliti atau pemikir dalam menyusun dan memahami informasi secara sistematis. mengumpulkan data dengan cara mengscrapping data pada google collab, Scraping data, Observasi metode pengumpulan data dengan mengamati peristiwa Pengumpulan data dalam pembuatan sistem ini dilakukan melalui studi literatur. Proses studi literatur meliputi pengolahan informasi yang relevan dengan pengembangan sistem serta pencatatan yang berasal dari berbagai bahan pustaka Hasil evaluasi menggunakan ROUGE menunjukkan bahwa metode ini efektif dalam mengekstraksi informasi penting dari ulasan, memungkinkan untuk menghasilkan ringkasan yang relevan dan informatif dengan nilai akurasi recall 93,78%.. memberikan hasil yang signifikan dalam melakukan. peringkasan teks pada ulasan aplikasi Tokopedia di Play Store. Dengan membangun graf berbobot dari matriks kesamaan kosinus berdasarkan representasi vektor TF-IDF.

Kata Kunci : Analisis Sentimen; Text Summarization; Algoritme Textrank; Ulasan.

Abstract

How does text summarization work in reviewing the Tokopedia application Google Play Store How to implement the text rank algorithm in doing so text summarization in the Tokopedia application review on Playstore. Measuring accuracy and the performance of the text rank algorithm in text summarization in the Tokopedia application review in google playstore How effective is the textrank algorithm when used for text summarization A framework is a structure used to help researchers or thinkers in compiling and understanding information systematically. collect data by scraping data on Google Collab, scraping data, Observation is a data collection method by observing events. The data collection method used in creating this system was literature study. Literature studies are carried out by processing information related to manufacturing system, making notes sourced from library materials. Evaluation results using ROUGE shows that this method is effective in extracting important information from the review, allowing to produce a summary that relevant and informative with a recall accuracy value of 93.78%. results in summarizing text in reviews Tokopedia application on the Play Store. By building a weighted graph from matrices cosine similarity based on TF-IDF vector representation.

Keywords : Sentiment Analysis; Text Summarization; Algorithm Textrank; Review.

1. PENDAHULUAN

Saat ini, persaingan bisnis di Indonesia semakin intens dan ketat. Hal ini terlihat dari munculnya banyak usaha baru. Kemajuan teknologi dan kemudahan akses internet telah memungkinkan para pelaku bisnis untuk beralih dari metode manual (Soviyana et al., 2022) Di Implementasi Text Summarization Review Aplikasi Tokopedia Di Google Playstore Menggunakan Algoritma Text Rank

era modern ini, Internet bisa diakses dengan mudah kapan pun dan di mana pun. Kemudahan serta jangkauan luas ini memungkinkan banyak orang untuk tetap terhubung tanpa terbatas oleh waktu maupun tempat. Selain sebagai alat komunikasi dan sumber informasi, internet juga telah berkembang menjadi platform untuk berbagai aktivitas ekonomi. Aktivitas perdagangan yang dahulu berlangsung secara langsung antara penjual dan pembeli, kini dapat dilakukan secara online, mempermudah proses perdagangan dan membuka peluang baru bagi banyak pelaku bisnis (Barkatullah, 2017) Kehadiran internet berperan penting dalam meningkatkan efektivitas dan efisiensi berbagai kegiatan bisnis. Dengan adanya internet, para pelaku bisnis dapat menjalankan operasi mereka dengan lebih mudah dan cepat. Internet menyediakan berbagai fasilitas yang memungkinkan kegiatan bisnis dilakukan dengan lebih lancar, mulai dari komunikasi dengan pelanggan hingga pengelolaan transaksi dan logistik. Kondisi ini menjadikan internet sebagai sarana yang sangat vital dalam bisnis modern, karena mampu membantu perusahaan meraih tujuan mereka secara lebih efisien dan produktif. (Setiawan, 2017). Di Indonesia, jumlah pengguna internet terus mengalami peningkatan setiap tahunnya. Berdasarkan laporan terbaru dari Hootsuite dan We Are Social, hingga Januari 2023 jumlah pengguna internet di Indonesia telah mencapai 212,9 juta orang. Pertumbuhan e-commerce juga memberikan dampak signifikan bagi para pelaku bisnis, karena memungkinkan mereka untuk memperluas cakupan pasar secara lebih luas (Avriyanti, 2020).

2. METODE / ALGORITMA

Struktur atau panduan yang dimanfaatkan oleh peneliti atau pemikir sebagai alat untuk merancang dan memahami ide serta informasi secara sistematis. Panduan ini berfungsi sebagai kerangka konseptual yang membantu dalam merumuskan dan mengorganisir argumen, teori, atau hipotesis yang akan dibahas atau dianalisis dalam suatu penelitian atau proyek.

Metode Pengumpulan Data

Salah satu teknik pengumpulan data yang digunakan penulis dalam pengembangan sistem ini adalah studi literatur. Metode ini dilakukan dengan membaca, mengutip, dan mengolah berbagai informasi yang berhubungan dengan proses pembuatan sistem, serta mencatat referensi dari sumber pustaka yang relevan dan mendukung pembangunan sistem tersebut.

Penulis mengumpulkan dan mempelajari berbagai sumber berupa data serta definisi informasi yang berasal dari sejumlah jurnal dan buku yang relevan dengan penelitian serta pengembangan sistem ini, juga dari beberapa situs online yang kredibel. Seluruh referensi tersebut digunakan untuk mendukung penyusunan bab pendahuluan, landasan teori, dan metode penelitian

Pengumpulan Data

Pengumpulan data dapat dilakukan melalui berbagai metode, termasuk teknik kualitatif seperti observasi, wawancara, dan studi kasus, serta teknik kuantitatif seperti eksperimen, survei, dan penggunaan data sekunder.

Dalam pengumpulan data, peneliti harus memperhatikan beberapa aspek penting seperti definisi variabel yang jelas, pengembangan instrumen pengumpulan data yang efektif, dan penerapan prosedur pengumpulan yang konsisten.

Data yang terkumpul kemudian dianalisis untuk menemukan pola, hubungan, atau temuan yang dapat memberikan wawasan atau menjawab pertanyaan penelitian. Pengumpulan data yang baik memerlukan perencanaan yang matang dan penerapan metode yang sesuai untuk

memastikan bahwa data yang diperoleh representatif dan relevan dengan hipotesis atau tujuan penelitian yang telah ditetapkan.

Penulis mengumpulkan data dengan cara mengscraping data *pada google collab*, Scraping data (atau *web scraping*) adalah proses otomatis untuk mengekstrak informasi atau data dari halaman web. Teknik ini digunakan untuk mengambil data yang terstruktur atau tidak terstruktur dari berbagai situs web dan menyimpannya dalam format yang dapat digunakan untuk analisis lebih lanjut, visualisasi, atau aplikasi lainnya.

3. HASIL DAN PEMBAHASAN

Bagian hasil dan pembahasan merupakan segmen dalam laporan penelitian yang memaparkan temuan observasi atau eksperimen, disertai analisis, interpretasi data, serta perbandingan dengan teori atau penelitian sebelumnya

Scrapping Data

Pengumpulan Data: Web scraping memungkinkan pengguna untuk mengumpulkan data dari berbagai sumber online secara otomatis. Ini dapat mencakup data seperti teks, gambar, harga produk, ulasan pelanggan, dan banyak lagi.

Automatisasi: Scraping data dapat dilakukan secara otomatis menggunakan perangkat lunak atau skrip yang diprogram untuk mengunjungi halaman web, mengekstrak data, dan menyimpannya dalam format yang diinginkan tanpa interaksi manusia yang intensif.

Teknik dan Alat: Ada berbagai teknik dan alat yang digunakan untuk melakukan web scraping, termasuk penggunaan perpustakaan Python seperti BeautifulSoup atau Scrapy, serta penggunaan alat-alat khusus yang dapat disesuaikan untuk tujuan scraping tertentu

Hukum dan Etika: Praktik web scraping harus memperhatikan hukum-hukum yang berlaku, terutama mengenai hak cipta dan kebijakan privasi. Beberapa situs web memiliki ketentuan yang melarang penggunaan scraping data mereka tanpa izin

Penerapan: Scraping data digunakan dalam berbagai aplikasi seperti pemantauan harga produk, analisis sentimen media sosial, penelitian pasar, dan banyak lagi. Ini memberikan akses cepat dan efisien ke informasi yang diperlukan dari internet

Preprocessing

Case folding adalah proses di mana semua perbedaan kapitalisasi atau huruf besar dalam sebuah string yang sensitif terhadap kasus dihapus, sehingga semua huruf diubah menjadi huruf kecil. Proses ini sering dilakukan untuk memastikan konsistensi dalam pemrosesan teks, di mana perbedaan huruf besar atau kecil tidak lagi mempengaruhi hasil analisis atau perbandingan.

Ulasan	Hasil Case Folding
Kurir rekomendasi di ilangin aja Apalagi yg ngirimnya sap x Barang ga dateng2, bikin emosi STOP JUAL HEWAN DILINDUNGI	kurir rekomendasi di ilangin ajaapalagi yg ngirimnya sap x Barang ga dateng2, bikin emosi
Percuma ada promo kalo tidak bisa dipake	stop jual hewan dilindungi
	percuma ada promo kalo tidak bisa dipake

mantap	mantap
Cepet dan mantap	cepat dan mantap

Tabel 1. Hasil Case Folding

Filtering adalah proses menghapus tanda titik berulang, tanda baca, angka, spasi berlebihan, dan kata-kata yang tidak relevan (*stopword*).

Ulasan	Hasil Case Filtering
Kurir rekomendasi di ilangin aja Apalagi yg ngirimnya sap x Barang ga dateng2, bikin emosi	kurir rekomendasi di ilangin ajaapalagi yg ngirimnya sap x Barang ga dateng2, bikin emosi
STOP JUAL HEWAN DILINDUNGI	stop jual hewan dilindungi
Percuma ada promo kalo tidak bisa dipake	percuma ada promo kalo tidak bisa dipake
mantap	mantap
Cepet dan mantap	cepat dan mantap

Tabel 2. Hasil Case Filtering

Tokenizing atau Tokenisasi adalah proses membagi kalimat menjadi token-token individu

```
from sklearn.feature_extraction.text import TfidfVectorizer

# Inisialisasi TF-IDF Vectorizer
tfidf_vectorizer = TfidfVectorizer()

# Fit dan transform data konten menjadi vektor TF-IDF
tfidf_matrix = tfidf_vectorizer.fit_transform(data['content'])

print("Tokenizing:")
print(tfidf_matrix)
```

Gambar 1. Source Code Hasil TF-IDF

Representasi vektor adalah cara untuk mengubah data teks atau objek lainnya menjadi format numerik yang bisa diproses oleh komputer. Ini seperti mengubah teks menjadi angka sehingga komputer bisa memahaminya dan melakukan analisis atau perhitungan.

```
[11] # output representasi vektor TF-IDF
print("TF-IDF Matrix Shape:", tfidf_matrix.shape)

TF-IDF Matrix Shape: (100, 594)
```

Gambar 2. Menunjukkan Dimensi Dari TF-IDF Yang Telah Dihitung

Similarity Matrix Kode yang diberikan menggunakan pustaka scikit-learn untuk menghitung matriks kesamaan kosinus dari dua matriks TF-IDF (Term Frequency-Inverse Document Frequency).

```
from sklearn.metrics.pairwise import cosine_similarity

# Menghitung matriks kesamaan kosinus
cosine_similarities = cosine_similarity(tfidf_matrix, tfidf_matrix)

print("Matriks Kesamaan Kosinus:")
print(cosine_similarities)
```

Matriks Kesamaan Kosinus:

[1.	0.	0.	...	0.	0.14436658	0.]
[0.	1.	0.	...	0.10348067	0.	0.]
[0.	0.	1.	...	0.	0.	0.]
...							
[0.	0.10348067	0.	...	1.	0.	0.]
[0.14436658	0.	0.	...	0.	1.	0.05969799]
[0.	0.	0.	...	0.	0.05969799	1.]]

Gambar 3. Menghitung Matriks Kesamaan Kosinus Dari Matriks TF-IDF

Setelah mengetahui nilai kedekatan antar kata, tahap selanjutnya adalah merubah nilai similarity matrix menjadi node-node dalam graph, lalu diberikan bobot angka menggunakan page rank

```
import networkx as nx

# Membangun graf berbobot menggunakan NetworkX dari matriks kesamaan kosinus
G = nx.from_numpy_array(cosine_similarities)

print("Graph Berbobot:")
print(G)

# Menghitung skor TextRank
scores = nx.pagerank(G)
```

Graph Berbobot:
Graph with 100 nodes and 1187 edges

Gambar 4. Menghitung Skor Textrank.

4. KESIMPULAN

Dalam penelitian ini, kami berhasil menerapkan metode text summarization menggunakan algoritma TextRank untuk merangkum ulasan pengguna aplikasi Tokopedia dari Google Play Store. Proses ini melibatkan pengolahan teks untuk mengubah teks menjadi huruf kecil, penyaringan karakter non-alfanumerik, serta penggunaan TF-IDF untuk mewakili dan memahami makna teks. Hasil evaluasi menggunakan ROUGE menunjukkan bahwa metode ini efektif dalam mengekstraksi informasi penting dari ulasan, memungkinkan untuk menghasilkan ringkasan yang relevan dan informatif dengan nilai akurasi recall 93,78%. Dan Implementasi algoritma TextRank Penelitian ini menghasilkan temuan yang berarti dalam melaksanakan peringkasan teks pada ulasan aplikasi Tokopedia di Play Store. Dengan membangun graf berbobot dari matriks kesamaan kosinus berdasarkan representasi vektor TF-IDF, kami berhasil mengidentifikasi dan mengekstraksi kalimat-kalimat kunci yang mencerminkan esensi dari ulasan pengguna. Penggunaan skor TextRank untuk merangking kalimat-kalimat ini memberikan kemampuan untuk menghasilkan ringkasan teks yang tidak hanya singkat, tetapi juga mewakili inti dari setiap ulasan yang dianalisis.

5. REFERENSI

- [1] Avriyanti. (2020). *Http://jurnal.stiatabalong.ac.id/index.php/PubBis*.
- [2] Barkatullah, A. H. (2017). *Hukum Transaksi Elektronik Di Indonesia*.
- [3] Setiawan, A. B. (2017). Kebijakan Teknologi Informasi dan Komunikasi Untuk Mendorong Pembentukan Model Bisnis Masa Depan Policy of Information and Communication Technologies to Promote the Formation of Future Business Models. In *Jurnal Pekommas* (Vol. 2, Issue 2).
- [4] Soviyana, K., Kamila, M., Sugiyono, N., Adinugrah, H. H., Syariah, E., Ekonomi, F., Bisnis, D., Masuk, N., & Haryanto & Rudy. (2022). Strategi Pemasaran Umkm Produksi Tahu Desa Batang Dalam Menghadapi Persaingan Antar Pengusaha. *Journal of Social Work and Empowerment*, 2(1). <https://ejournal.catuspata.com/index.php/joswae>
- [5] Yusron, M. I., Bachtiar, F. A., & Fauzi, M. A. (2019). Peringkasan Review Konsumen Restoran Menggunakan Weighted Frequent Itemset Mining. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(5), 4961– 4970
- [6] Sulistyawati, U. S. (2024). Membangun Keunggulan Kompetitif melalui Platform E-Commerce: Studi Kasus Tokopedia. *Jurnal Manajemen Dan Teknologi*, 1(1), 43–56
- [7] Salim, S. (2022). Media Pelayanan Gereja di Era Digital. *NON MULTA SED MULTUM:(Bukan Jumlah Tetapi Mutu)*, 344.
- [8] Juniansyah, B. D., Susanto, E. R., & Wahyudi, A. D. (2020). Pembuatan E- Commerce Pemesanan Jasa Event Organizer Untuk Zero Seven Entertainment. *Jurnal Tekno Kompak*, 14(1), 41–46.
- [9] Bahtiar, R. A. (2020). Potensi, peran pemerintah, dan tantangan dalam pengembangan e-commerce di Indonesia [Potency, government role, and challenges of e-commerce development in Indonesia]. *Jurnal Ekonomi Dan Kebijakan Publik*, 11(1), 13–25
- [10] Avriyanti, S. (2020). Peran e-commerce untuk meningkatkan keunggulan kompetitif di era industri 4.0 (studi pada ukm yang terdaftar pada dinas koperasi, usaha kecil dan menengah kabupaten tabalong). *PubBis: Jurnal Pemikiran Dan Penelitian Administrasi Publik Dan Administrasi Bisnis*, 4(1), 83–99.
- [11] Amanah, D., & Layla, A. A. (2019). Pengaruh harga dan kelengkapan produk terhadap keputusan pembelian konsumen online shop tokopedia. com di fakultas ekonomi jurusan manajemen Universitas Negeri Medan. *Liabilities (Jurnal Pendidikan Akuntansi)*, 2(2), 125–138.
- [12] Zamzam, M. A., Crysdiyan, C., & Holle, K. F. H. (2020). Sistem automatic text summarization menggunakan algoritma textrank. *MATICS: Jurnal Ilmu Komputer Dan Teknologi Informasi (Journal of Computer Science and Information Technology)*, 12(2), 111–116.