



# PENERAPAN METODE *VECTOR SPACE MODEL* (VSM) DENGAN TF-IDF DAN *COSINE SIMILARITY* PADA SISTEM TEMU KEMBALI INFORMASI LOWONGAN PEKERJAAN

Fadli Maghfirli <sup>a,1,\*</sup>, M. Aditya Firmansyah <sup>b,2</sup>, Laili Cahyani <sup>c,3</sup>

<sup>a,b,c</sup> Universitas Trunojoyo Madura, Jalan Raya Telang, Kecamatan Kamal, Kabupaten Bangkalan, Jawa Timur 69162.

<sup>1</sup> fadlimaghfirli5@gmail.com\*; <sup>2</sup> muhamad.aditya0007@gmail.com; <sup>3</sup> laili.cahyani@trunojoyo.ac.id

\* corresponding author

## ARTICLE INFO

## ABSTRACT

### Keywords

*Information Retrieval, Job Vacancy, TF-IDF, Cosine Similarity, Vector Space Model.*

Finding relevant job vacancy information is often a challenge for job seekers due to the large volume of data on the internet and the limitations of exact keyword matching, which often produces inaccurate results. This study aims to build a job vacancy Information Retrieval (IR) system using the Vector Space Model (VSM) approach. The system applies Term Frequency-Inverse Document Frequency (TF-IDF) for term weighting and Cosine Similarity to measure the relevance between user queries and job documents. The study utilized a dataset of 350 documents, consisting of 300 valid job postings and 50 noise documents. The text preprocessing stages included case folding, tokenizing, filtering, and stemming. System performance was evaluated using Precision, Recall, and F-Measure metrics on 10 different search queries. The test results demonstrated high accuracy in retrieving relevant documents, achieving an average Precision of 0.840 (84%) at the top-10 threshold and 0.880 (88%) at the top-15 threshold. These results indicate that the combination of TF-IDF and Cosine Similarity is effective in filtering out irrelevant documents and ranking job vacancies according to the user's needs.

## 1. Pendahuluan

Perkembangan teknologi informasi yang pesat telah mengubah cara masyarakat dalam mencari pekerjaan. Internet menyediakan platform yang luas bagi penyedia kerja untuk mempublikasikan lowongan, namun hal ini memunculkan masalah baru berupa ledakan informasi (*information overload*). Pencari kerja seringkali kesulitan menemukan posisi yang sesuai dengan keahlian mereka karena banyaknya dokumen yang tidak relevan yang muncul pada hasil pencarian. Oleh karena itu, dibutuhkan sebuah sistem temu kembali informasi (*Information Retrieval*) yang mampu menyaring dan memeringkat dokumen berdasarkan tingkat relevansinya terhadap kebutuhan pengguna [1].

Permasalahan utama pada sistem pencarian lowongan kerja sederhana adalah penggunaan metode pencocokan string biasa yang kaku. Jika pencari kerja memasukkan kata kunci yang tidak persis sama dengan yang ada di database, sistem seringkali gagal menampilkan hasil yang relevan. Diperlukan pendekatan berbasis pembobotan kata dan pengukuran kemiripan vektor untuk mengatasi variasi bahasa dalam deskripsi pekerjaan [2].

Beberapa penelitian terdahulu telah menerapkan metode serupa untuk kasus yang berbeda. Algoritma TF-IDF dan *Cosine Similarity* pada sistem rekomendasi lowongan pekerjaan dan menghasilkan tingkat akurasi sebesar 80% [3]. Penelitian tersebut menunjukkan bahwa kombinasi kedua metode ini efektif dalam mencocokkan profil pelamar dengan deskripsi pekerjaan. Selain itu, dalam penelitian mengenai sistem pencarian pekerjaan menyebutkan bahwa meskipun metode TF-IDF membutuhkan waktu komputasi lebih lama dibandingkan *full query*, hasil pencarian yang diberikan jauh lebih relevan [4].

Efektivitas *Cosine Similarity* juga dikuatkan pada sistem temu kembali informasi pasal KUHP, di mana metode ini mampu mengukur kedekatan dokumen hukum dengan *queri* pengguna secara presisi [5]. Sementara itu, pendekatan *Content-Based Filtering* dengan algoritma yang sama untuk merekomendasikan pekerjaan berdasarkan keterampilan, yang membuktikan fleksibilitas algoritma ini

<https://doi.org/10.47111/JTI>

Available online at <https://e-journal.upr.ac.id/index.php/JTI>

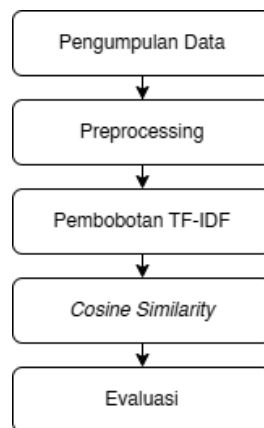


dalam domain ketenagakerjaan [6]. Namun, sebagian besar penelitian tersebut berfokus pada sistem rekomendasi (*push information*), sedangkan penelitian ini berfokus pada mekanisme pencarian aktif (*pull information*) melalui kueri pengguna pada sistem temu kembali informasi.

Berdasarkan latar belakang tersebut, penelitian ini mengusulkan penerapan metode Term *Frequency-Inverse Document Frequency* (TF-IDF) untuk pembobotan kata dan *Cosine Similarity* untuk pengukuran relevansi pada sistem pencarian lowongan pekerjaan. TF-IDF dipilih karena kemampuannya memberikan bobot tinggi pada kata unik yang mencirikan suatu dokumen [7], sedangkan *Cosine Similarity* dipilih karena efektivitasnya dalam mengukur kemiripan dua dokumen teks yang dinormalisasi, sehingga tidak terpengaruh oleh panjang pendeknya dokumen. Tujuan dari penelitian ini adalah membangun sistem yang dapat memudahkan pencari kerja mendapatkan informasi lowongan yang paling relevan dengan kata kunci yang mereka masukkan.

## 2. Metodologi Penelitian

Metodologi penelitian ini menjelaskan tahapan-tahapan yang dilakukan dalam membangun sistem temu kembali informasi (*Information Retrieval*) untuk pencarian lowongan pekerjaan. Penelitian ini menerapkan model *Vector Space Model* (VSM) dengan pembobotan TF-IDF dan pengukuran kemiripan menggunakan *Cosine Similarity*. Secara umum, alur penelitian terdiri dari pengumpulan data, *preprocessing*, pembobotan, perhitungan kemiripan, dan evaluasi [8].



Gambar 1. Tahapan Penelitian

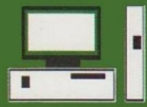
### 2.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang terdiri dari dua kategori utama untuk menguji performa sistem dalam membedakan dokumen relevan dan tidak relevan. Total dataset berjumlah 350 dokumen.

### 2.2. *Preprocessing*

*Preprocessing* adalah proses text mining yang digunakan untuk mendapatkan informasi berkualitas tinggi dari teks dan merupakan tahapan awal untuk menyiapkan teks menjadi data yang akan diolah lebih lanjut [9]. Karena algoritma pencarian tidak dapat mengubah teks secara langsung dalam proses, proses *preprocessing* diperlukan untuk mengubah teks menjadi data numerik [10]. Tahapan *preprocessing* yang dilakukan adalah sebagai berikut [11]:

1. *Case Folding*: Mengubah seluruh huruf dalam dokumen menjadi huruf kecil (*lowercase*) untuk menyeragamkan format teks.
2. *Cleaning*: Menghapus karakter angka, tanda baca, dan karakter non-alfabet menggunakan fungsi `re.sub(r'^a-zs', '', text)` agar fokus hanya pada kata-kata yang bermakna.



3. *Stopword Removal*: Menghilangkan kata-kata umum yang sering muncul namun tidak memiliki makna penting dalam pencarian (seperti "dan", "yang", "di"). Proses ini menggunakan daftar *stopword* dari *library* Sastrawi (*StopWordRemoverFactory*).
4. *Stemming*: Mengubah kata berimbuhan menjadi kata dasarnya (contoh: "membutuhkan" menjadi "butuh") menggunakan *StemmerFactory* dari Sastrawi. Langkah ini penting untuk mengatasi variasi morfologi kata dalam Bahasa Indonesia.

### 2.3. Pembobotan TF-IDF

Setelah data bersih, tahap selanjutnya adalah pembobotan kata menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini mengubah data teks menjadi representasi vektor numerik (*matrix*) [12].

### 2.4. Cosine Similarity

Nilai *Cosine Similarity* berkisar antara 0 hingga 1. Nilai mendekati 1 menunjukkan kemiripan dokumen yang tinggi dengan *query*, sedangkan nilai 0 menunjukkan ketidakmiripan sama sekali. Rumus perhitungan yang digunakan adalah:

### 2.5. Evaluasi

Evaluasi dilakukan untuk mengukur performa sistem dalam menampilkan dokumen yang relevan. Pengujian dilakukan menggunakan 10 *query* pencarian yang berbeda terkait lowongan pekerjaan (misalnya: "IT", "sales admin").

## 3. Hasil dan Pembahasan

### 3.1. Pengumpulan Data

Pada tahap implementasi, data yang telah dikumpulkan berhasil dimuat ke dalam sistem menggunakan *library* Pandas. Dataset yang diproses terdiri dari 350 baris data dengan tiga atribut utama yang digunakan dalam penelitian, yaitu judul lowongan (*job\_title*), deskripsi pekerjaan (*job\_description*), dan label kelas (Relevansi).

Table 1. Tabel Dataset

relevansi	job_title	job_description
Relevan	Admin Online Shop	Deskripsi Pekerjaan:Dapat menggunakan aplikasi E-Commerce/Market Place (contoh: Tokopedia, Blibli, Shopee, dan lain-lain) Bertanggung jawab terhadap penjualan online Memasarkan produk broadcast chat dan social media. Memonitor penjualan dan pengiriman untuk orderan online Membantu menyelesaikan jika ada orderan online yang bermasalah Menjawab chat / pertanyaan konsumen online Memberikan pelayanan yang terbaik bagi konsumen Selalu belajar hal baru & mengikuti perkembangan aplikasi-aplikasi marketplace dan sosmed. Kualifikasi: Maksimal Usia 23 tahun Lulusan SMASMK/ Sederajat Tidak sedang kullah Familiar dengan homputer, sosmed FB, IG, marketplace, dan MS Office Amanah, disiplin dan bertanggung jawab Kreatif dan mau belajar Diutamakan domisili Serpong dan sekitarnya
Relevan	Procurement Department Head	Job Role: 1. Responsible for material availability for production heavy equipment and supporting equipment to ensure support to operational needs, with proper specification, budget, quantity and time required. 2. Ensure complete database of suppliers/vendors/subcontractors, contract management, material delivery, scrapeselling activities from Jobsite accordingly to plan and procedure. 3. Ensuring subcontractor requirements for heavy equipment/units in Production, Engineering & Plant Requirement(s) Education S1 Mechanical Engineering / Industrial Engineering / Economics Work Experience(s) 5 Years in the same position, preferably from mining services Industry Preference: Mining Services Technical

		Skill: Heavy Equipment Knowledge, Negotiation Skill, Contract Management Good command in English is a must
...	...	...
Non-Relevan	Sales Asuransi Jiwa Konvensional Tergeser Platform Insurtech	Generasi muda yang lebih memilih membeli produk asuransi mikro melalui aplikasi (Insurtech) tanpa perantara agen membuat perusahaan asuransi konvensional melakukan penyesuaian. Ribuan tenaga pemasar asuransi (insurance agents) dengan kinerja di bawah rata-rata diputus kontrak kemitraannya karena perusahaan menutup kantor-kantor keagenan fisik yang sepi peminat demi efisiensi biaya sewa.
Non-Relevan	Rumah Sakit Swasta Lakukan Efisiensi Staf Admin Pendaftaran	Digitalisasi layanan kesehatan melalui aplikasi pendaftaran online dan anjungan mandiri di rumah sakit swasta mengurangi antrean fisik. Dampaknya, manajemen rumah sakit mulai mengurangi jumlah staf admin pendaftaran dan rekam medis manual, mengalihkan fokus anggaran pada tenaga medis profesional dibandingkan tenaga administratif.

### 3.2. Preprocessing

Setelah data dikumpulkan, tahapan selanjutnya adalah *preprocessing*. Tahap ini bertujuan untuk membersihkan dan mempersiapkan data teks agar optimal saat diproses oleh algoritma. Mengingat data yang digunakan berbahasa Indonesia, proses ini memanfaatkan pustaka Sastrawi untuk *stemming* dan *stopword removal*, serta *Regular Expression* (Regex) untuk pembersihan karakter.

Dan berikut hasil *preprocessing* dataset, dapat dilihat pada Gambar 2. Hasil Preprocessing

	job_title	job_description	Relevansi	clean_text
0	Facility Maintenance & Smart Warehouse Manager	Deskripsi PekerjaanRequirements :D3/S1 from reputable University Majoring Electrical/Mechanical ...	Relevan	deskripsi pekerjaanrequirements ds1 from reputable university majoring electricalmechanical engi...
1	Procurement Department Head	Job Role: 1. Responsible for material availability for production heavy equipment and supporting ...	Relevan	job role responsible for material availability for production heavy equipment and supporting equ...
2	SALES ADMIN	Internal Sales & AdminJob Description :We are looking for a qualified Internal Sales & Admin to ...	Relevan	internal sales adminjob description we are looking for a qualified internal sales admin to join ...
3	City Operation Lead Shopee Express (Cirebon)	Job Description Responsible for HSE implementation and ensure people safety within assigned City...	Relevan	job descriptionresponsible for hse implementation and ensure people safety within assigned city...
4	Japanese Interpreter	Overview: Our clients is manufacture for automotive parts which located in MM2100, Bekasi.Postil...	Relevan	overview our clients is manufacture for automotive parts which located in mm bekasiposition japa...
...	...	...	...	...
345	Manajer HRD Terdampak Kebijakan "Hiring Freeze" Jangka Panjang	Banyak perusahaan teknologi dan startup yang menerapkan kebijakan pembekuan rekrutmen (hiring fr...	Non-Relevan	banyak usaha teknologi startup terap beku rekrutmen hiring freeze panjang tahun akhir putu...
346	Manajer Operasional Logistik Dikurangi Pasca-Optimasi Rute AI	Perusahaan logistik besar mulai menerapkan sistem manajemen armada berbasis Artificial Intelligence...	Non-Relevan	usaha logistik besar mulai terap sistem manajemen armada bas artificial intelligence mampu atur ...
347	Medical Representative" Farmasi Dikurangi Akibat Pengadaan Digital RS	Transformasi digital di sektor kesehatan membuat rumah sakit kini memesan obat-obatan melalui st...	Non-Relevan	transformasi digital sektor sehat buat rumah sakit kini mes obatobatan lalu sistem ecatalogue pu...
348	Sales Asuransi Jiwa Konvensional Tergeser Platform Insurtech	Generasi muda yang lebih memilih membeli produk asuransi mikro melalui aplikasi (Insurtech) tanp...	Non-Relevan	generasi muda lebih pilih beli produk asuransi mikro lalu aplikasi insurtech antara agen buat us...
349	Rumah Sakit Swasta Lakukan Efisiensi Staf Admin Pendaftaran	Digitalisasi layanan kesehatan melalui aplikasi pendaftaran online dan anjungan mandiri di rumah...	Non-Relevan	digitalisasi layan sehat lalu aplikasi daftar online anjung mandiri rumah sakit swasta kurang an...

350 rows x 4 columns

Gambar 2. Hasil Preprocessing

### 3.3 Pembobotan TF-IDF

Setelah data teks bersih, dilakukan proses pembobotan kata menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini mengubah data teks menjadi matriks angka yang merepresentasikan bobot pentingnya sebuah kata dalam dokumen terhadap keseluruhan korpus data.

Proses ini diimplementasikan menggunakan modul `TfidfVectorizer` dari pustaka Scikit-Learn. Berikut kode sumber pembobotan TF-IDF:

```
# Inialisasi TfidfVectorizer
vectorizer = TfidfVectorizer()

# Transformasi data teks menjadi matriks bobot TF-IDF
tfidf_matrix = vectorizer.fit_transform(df_final['clean_text'])
```

### 3.4 Cosine Similarity

Tahap inti dari sistem temu kembali informasi ini adalah perhitungan kemiripan antara *query* (kata kunci pencarian) yang dimasukkan pengguna dengan dokumen lowongan kerja yang tersedia. Metode *Cosine Similarity* menggunakan kata kunci (*keyword*) dari dokumen untuk mengukur tingkat kesamaan



(*similarity*) antara dua objek yang diwakili dalam dua vektor [7]. Berikut kode sumber untuk *Cosine Similarity*:

```
def search_engine(query, top_n=15):
    # Preprocess query user
    clean_query = preprocessing(query)

    # Transform ke vector
    query_vec = vectorizer.transform([clean_query])

    # Hitung kemiripan (Cosine Similarity)
    similarity_scores = cosine_similarity(query_vec,
    tfidf_matrix).flatten()

    # Urutkan index dari score tertinggi
    top_indices = similarity_scores.argsort()[::-1][:top_n]

    # ... (kode untuk menampilkan hasil) ...
```

Berdasarkan uji coba dengan query "IT", sistem berhasil menampilkan dokumen yang memiliki skor relevansi tertinggi, seperti terlihat pada Gambar 3. Hasil Pencarian Query IT



Gambar 3. Hasil Pencarian Query IT

### 3.5 Evaluasi

Dalam konteks sistem yang telah dibuat ini, evaluasi bertujuan untuk menentukan seberapa jauh hasil rekomendasi sistem memenuhi harapan atau preferensi pencari lowongan kerja [6]. Proses evaluasi mengacu pada metode pengujian yang dilakukan dalam penelitian [5], yaitu dengan menggunakan perhitungan Precision, Recall, dan F-Measure.

Pengujian dilakukan dengan memasukkan 10 (sepuluh) kata kunci (*query*) yang berbeda ke dalam sistem. Kata kunci ini dipilih untuk mewakili berbagai bidang pekerjaan yang terdapat dalam *dataset*. Setiap *query* akan diproses oleh sistem, dan sistem akan menampilkan 10 dokumen teratas (Top-10) yang memiliki nilai *Cosine Similarity* tertinggi.

Daftar kata kunci yang digunakan dalam skenario pengujian dapat dilihat pada Table 2. Query pencarian

Table 2. Query pencarian

Query pencarian
IT
Manager
Sales



Bank
Staff
Teknik
Admin
Marketing
Supervisor
Design

### 3.5.1 Hasil Pengujian

Hasil pencarian dokumen dengan *query* IT

Table 3. Hasil Pencarian Query IT

Threshold	job_title	Relevansi
10	IT Governance Supervisor	Relevan
	Staf IT Helpdesk Level 1 Digantikan Chatbot Cerdas	Non-Relevan
	Badai PHK Tech Winter Masih Berlanjut di Tahun 2025	Non-Relevan
	IT SECURITY COORDINATOR	Relevan
	IT Support	Relevan
	SPV IT	Relevan
	IT Project Management	Relevan
	Odo Developer / IT Full Stack Developer	Relevan
	Startup "Unicorn" Indonesia Lakukan Efisiensi Tim IT dan Produk	Non-Relevan
	Sales B2B Perangkat Lunak (SaaS) Dikurangi Akibat Efisiensi Klien	Non-Relevan
15	SECURITY ENGINEER	Relevan
	IT Staff	Relevan
	IT Software Developer Staff	Relevan
	IT Support Staff	Relevan
	Backend Developer	Relevan

Precision = 0.73  
Recal = 0.037  
F-measure = 0.070

Kemudian Query 2 sampai Query 10 dilakukan perhitungan seperti pada Query 1 di atas untuk mencari nilai precision, recall dan f.measure pada threshold 5, 10 dan 15. Hasil dari pencarian precision, recall dan f.measure pada semua query dapat dilihat pada table berikut:

Table 4. Hasil Evaluasi Precision, Recall, F-measure pada Threshold 10

	Precision	Recall	F-measure
Query 1	0.600	0.020	0.039
Query 2	0.800	0.027	0.052
Query 3	0.900	0.030	0.058
Query 4	0.700	0.023	0.045
Query 5	0.800	0.027	0.052
Query 6	1.000	0.033	0.065
Query 7	0.800	0.027	0.052

<https://doi.org/10.47111/JTI>

Available online at <https://e-journal.upr.ac.id/index.php/JTI>



Query 8	1.000	0.033	0.065
Query 9	0.800	0.027	0.052
Query 10	1.000	0.033	0.065
Rata-rata	0.840	0.028	0.054

Table 5. Hasil Evaluasi Precision, Recall, F-measure pada Threshold 15

	Precision	Recall	F-measure
Query 1	0.733	0.037	0.070
Query 2	0.867	0.043	0.083
Query 3	0.867	0.043	0.083
Query 4	0.800	0.040	0.076
Query 5	0.867	0.043	0.083
Query 6	0.933	0.047	0.089
Query 7	0.867	0.043	0.083
Query 8	1.000	0.050	0.095
Query 9	0.867	0.043	0.083
Query 10	1.000	0.050	0.095
Rata-rata	0.880	0.044	0.084

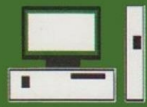
## 5. Kesimpulan

Berdasarkan implementasi dan hasil pengujian yang telah dilakukan pada sistem temu kembali informasi lowongan pekerjaan menggunakan metode *Vector Space Model* (VSM) dengan pembobotan TF-IDF dan *Cosine Similarity*, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Efektivitas Metode: Penerapan kombinasi metode pembobotan TF-IDF dan pengukuran kemiripan *Cosine Similarity* terbukti efektif dalam menyaring dan memeringkat dokumen lowongan pekerjaan. Sistem mampu membedakan antara dokumen lowongan kerja yang relevan dan dokumen *noise* (berita atau artikel umum) dengan baik.
2. Akurasi Pencarian (*Precision*): Berdasarkan pengujian terhadap 350 dokumen (300 data relevan dan 50 data *noise*) menggunakan 10 *query* uji yang berbeda, sistem menunjukkan tingkat relevansi yang tinggi pada hasil pencarian teratas. Rata-rata nilai *Precision* yang diperoleh adalah 0,840 (84%) pada batasan hasil pencarian (*threshold*) 10 dokumen teratas, dan meningkat menjadi 0,880 (88%) pada *threshold* 15 dokumen teratas.
3. Kinerja pada Top-N: Hasil evaluasi menunjukkan bahwa semakin banyak dokumen yang ditampilkan (dari top-10 ke top-15), nilai presisi cenderung stabil bahkan meningkat, yang mengindikasikan bahwa dokumen-dokumen relevan terdistribusi dengan baik pada peringkat atas hasil pencarian sistem. Hal ini membuktikan bahwa sistem dapat memenuhi kebutuhan pengguna dalam mendapatkan informasi yang paling relevan secara cepat tanpa harus memilah banyak dokumen yang tidak sesuai.

## Daftar Pustaka

- [1] F. Sanjaya, "Pemanfaatan Sistem Temu Kembali Informasi dalam Pencarian Dokumen Menggunakan Metode Vector Space Model," *J. Inf. Technol.*, vol. 5, pp. 147–153, 2017.
- [2] K. R. Ula and M. Faisal, "Analisis Kesesuaian Lowongan Pekerjaan dan Kemampuan Pencari Kerja dengan Menggunakan Metode Cosine Similarity," vol. 24, no. 3, pp. 617–631, 2025.
- [3] V. M. Hersianty *et al.*, "Penerapan Algoritma TF-IDF dan Cosine Similarity dalam Sistem Rekomendasi Lowongan Pekerjaan," vol. 9, no. 1, pp. 1619–1625, 2025.
- [4] A. Tirtana, A. Zulkarnain, and Y. D. Listio, "Pembuatan Sistem Pencarian Pekerjaan," vol. 13, no. 2, pp. 91–100, 2019.



- [5] S. Yusuf, M. A. Fauzi, and K. C. Brata, "Sistem Temu Kembali Informasi Pasal-Pasal KUHP ( Kitab Undang- Undang Hukum Pidana ) Berbasis Android Menggunakan Metode Synonym Recognition dan Cosine Similarity," vol. 2, no. 2, 2018.
- [6] A. P. Rachmayantie, N. Marliza, and A. Irianto Sukowati, "Sistem Rekomendasi Pekerjaan Menggunakan Algoritma TF-IDF dan Cosine Similarity," vol. 9, 2025.
- [7] A. Apriani, H. Zakiyudin, and K. Marzuki, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta Application of the Cosine Similarity Algorithm and Weighting of the TF-IDF System for New Student Admissions on Private Campuses," vol. 3, no. 1, pp. 19–27, 2021, doi: 10.30812/bite.v3i1.1110.
- [8] D. B. Elnursa, V. Nofriana, A. Syamsuri, and L. Cahyani, "Sistem Rekomendasi Pemilihan Program MSIB Bagi Mahasiswa Pendidikan Informatika," vol. 2, pp. 34–45, 2023.
- [9] D. Agatha, K. Kusriani, and E. T. Luthfi, "Analisis Pembobotan Kata Pada Klasifikasi Text Mining," vol. 3, no. 2, pp. 179–184, 2019.
- [10] R. Melita, V. Amrizal, H. B. Suseno, and T. Dirjam, "Penerapan Metode Term Frequency Inverse Document Frequency (TF-IDF) dan Cosine Similarity Pada Sistem Temu Kembali Informasi Untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Syarah Umdatil Ahkam)," vol. 11, no. 2, 2018.
- [11] D. Harjanto, S. N. Endah, and N. Bahtiar, "Sistem Temu Kembali Informasi pada Dokumen Teks Menggunakan Metode Term Frequency Inverse Document Frequency (TF-IDF).pdf," 2012.
- [12] J. E. B. Sinulingga and H. C. K. Sitorus, "Analisis Sentimen Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF Sentiment Analysis of Public towards Indonesian Horror Films Using SVM and TF- IDF Methods," vol. 14, no. April, pp. 42–53, 2024.