

**SISTEM REKOMENDASI JURNAL ILMIAH TEKNOLOGI INFORMASI
TERINDEKS SINTA BERBASIS CONTENT-BASED FILTERING
DENGAN *EMBEDDING* LABSE**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh:

**MUHAMMAD RIFQI NAUFAL IRSYAD
09010622010**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA
2026**

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini,

Nama : Muhammad Rifqi Naufal Irsyad
NIM : 09010622010
Program Studi : Sistem Informasi
Angkatan : 2022

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul: “SISTEM REKOMENDASI JURNAL ILMIAH TEKNOLOGI INFORMASI TERINDEKS SINTA BERBASIS CONTENT-BASED FILTERING DENGAN *EMBEDDING* LABSE”. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 05 Mei 2026

Yang menyatakan,



(Muhammad Rifqi Naufal Irsyad)

09010622010

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi Oleh:

NAMA : MUHAMMAD RIFQI NAUFAL IRSYAD

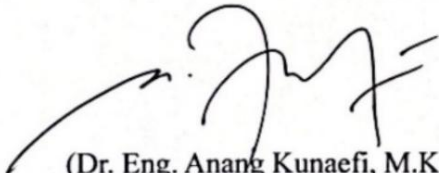
NIM : 09010622010

JUDUL : SISTEM REKOMENDASI JURNAL ILMIAH TEKNOLOGI INFORMASI
TERINDEKS SINTA BERBASIS CONTENT-BASED FILTERING DENGAN
EMBEDDING LABSE

Ini telah diperiksa dan disetujui untuk diujikan.

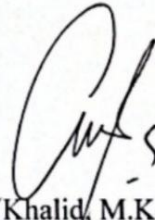
Surabaya, 13 Mei 2026

Dosen Pembimbing 1



(Dr. Eng. Anang Kunaefi, M.Kom)
NIP. 197911132014031001

Dosen Pembimbing 2



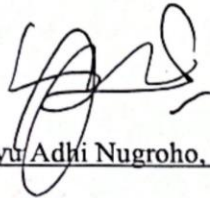
(Khalid, M.Kom)
NIP. 197906092014031002

LEMBAR PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi Muhammad Rifqi Naufal Irsyad ini telah dipertahankan
di depan tim penguji Skripsi
di Surabaya, 13 Mei 2026

Mengesahkan,
Dewan Penguji

Penguji I



(Bayu Adhi Nugroho, Ph.D.)

NIP. 197905182014031001

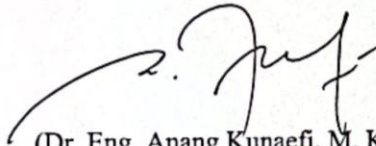
Penguji II



(Dwi Rolliawati, MT)

NIP. 197909272014032001

Penguji III



(Dr. Eng. Anang Kunaefi, M. Kom)

NIP. 197911132014031001

Penguji IV



(Khalid, M. Kom)

NIP. 197906092014031002

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya



(Dr. A. Saepul Hamdani, M.Pd.)

NIP. 196507312000031002

LEMBAR PERSETUJUAN PUBLIKASI



**KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN**

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Muhammad Rifqi Naufal Irsyad
NIM : 09010622010
Fakultas/Jurusan : Sains dan Teknologi/Sistem Informasi
E-mail address : rifqii.naufall37@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Skripsi Tesis Desertasi Lain-lain (.....)
yang berjudul :

SISTEM REKOMENDASI JURNAL ILMIAH TEKNOLOGI INFORMASI TERINDEKS SINTA BERBASIS

CONTENT-BASED FILTERING DENGAN EMBEDDING LABSE

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 20 Mei 2026

Penulis



Mubammad Rifqi Naufal Irsyad

ABSTRAK

SISTEM REKOMENDASI JURNAL ILMIAH TEKNOLOGI INFORMASI TERINDEKS SINTA BERBASIS CONTENT-BASED FILTERING DENGAN *EMBEDDING* LABSE

Oleh:

Muhammad Rifqi Naufal Irsyad

Peningkatan jumlah publikasi ilmiah di Indonesia menimbulkan tantangan baru bagi peneliti dalam memilih jurnal yang tepat untuk naskah mereka. Sistem pencarian jurnal yang ada saat ini umumnya masih mengandalkan pencocokan kata kunci secara leksikal, sehingga tidak mampu menangkap kesamaan makna antar teks yang menggunakan istilah berbeda. Penelitian ini mengembangkan sistem rekomendasi jurnal berbasis *content-based filtering* dengan memanfaatkan model *Language-agnostic BERT Sentence Embedding* (LaBSE) untuk merepresentasikan kesamaan semantik antara artikel dan karakteristik jurnal teknologi informasi terindeks SINTA. Data artikel berupa judul dan abstrak diperoleh dari portal Garba Rujukan Digital (GARUDA) pada rentang tahun 2022–2024. Evaluasi dilakukan menggunakan metrik *journal accuracy* dan *Mean Reciprocal Rank* (MRR). Hasil eksperimen menunjukkan bahwa kombinasi judul dan abstrak menghasilkan performa terbaik dengan nilai Acc@30 sebesar 0,836 dan MRR sebesar 0,244. Model LaBSE juga mengungguli model embedding lain yang diuji, termasuk BERT multilingual (MRR 0,213), XLM-RoBERTa (MRR 0,211), USE (MRR 0,201), SBERT (MRR 0,205), SPECTER (MRR 0,172), dan SciBERT (MRR 0,081). Keunggulan LaBSE ini disebabkan oleh arsitektur *dual-encoder*-nya yang dioptimalkan untuk *cross-lingual retrieval*, sehingga mampu menjembatani representasi semantik teks berbahasa Indonesia dan Inggris yang umum ditemukan dalam metadata jurnal nasional. Penelitian ini membuktikan bahwa pendekatan semantik berbasis LaBSE dapat menjadi solusi efektif dalam membantu peneliti menemukan jurnal publikasi yang relevan secara lebih cepat dan akurat.

Kata kunci: sistem rekomendasi jurnal, *content-based filtering*, LaBSE, SINTA, GARUDA

ABSTRACT

SINTA-INDEXED INFORMATION TECHNOLOGY JOURNAL RECOMMENDATION SYSTEM BASED ON CONTENT-BASED FILTERING WITH LABSE EMBEDDING

By:

Muhammad Rifqi Naufal Irsyad

The growing volume of scientific publications in Indonesia poses a significant challenge for researchers in selecting appropriate journals for their manuscripts. Existing journal search systems predominantly rely on lexical keyword matching, rendering them incapable of capturing semantic similarity between texts that use different terminology. This study develops a journal recommendation system based on content-based filtering by leveraging the Language-agnostic BERT Sentence Embedding (LaBSE) model to represent semantic similarity between articles and the characteristics of SINTA-indexed information technology journals. Article data, consisting of titles and abstracts, were collected from the Garba Rujukan Digital (GARUDA) portal spanning the years 2022–2024. Evaluation was conducted using journal accuracy and Mean Reciprocal Rank (MRR) metrics. Experimental results demonstrate that the combination of title and abstract achieves the best overall performance, with Acc@30 of 0.836 and MRR of 0.244. LaBSE also outperformed all other embedding models tested, including BERT multilingual (MRR 0.213), XLM-RoBERTa (MRR 0.211), USE (MRR 0.201), SBERT (MRR 0.205), SPECTER (MRR 0.172), and SciBERT (MRR 0.081). This superiority is attributed to LaBSE's dual-encoder architecture, which is explicitly optimized for cross-lingual retrieval, enabling it to bridge semantic representations across Indonesian and English texts commonly found in national journal metadata. This research demonstrates that a LaBSE-based semantic approach is an effective solution for helping researchers identify relevant publication journals more efficiently and accurately.

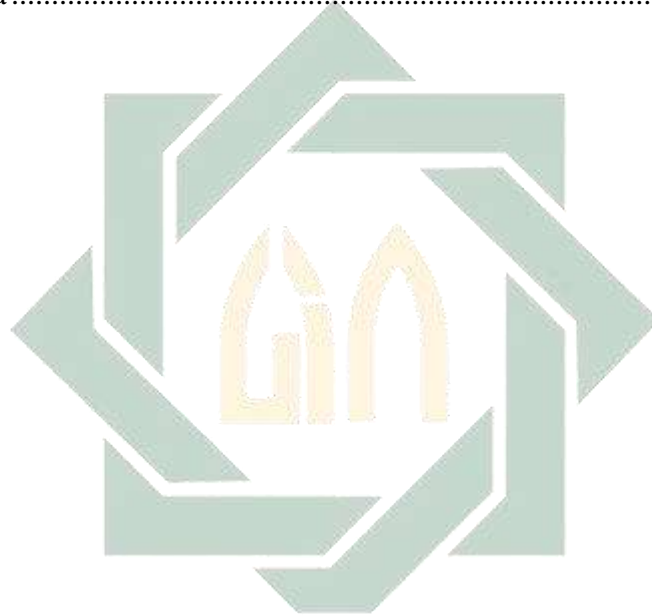
Keywords: *journal recommendation system, content-based filtering, LaBSE, SINTA, GARUDA*

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN TIM PENGUJI SKRIPSI	ii
LEMBAR PERNYATAAN KEASLIAN SKRIPSI	iii
LEMBAR PERSETUJUAN PUBLIKASI.....	iv
MOTTO.....	v
KATA PENGANTAR.....	vi
ABSTRAK	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	6
1.5.1 Akademik	6
1.5.2 Aplikatif	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Tinjauan Penelitian Terdahulu.....	7
2.2 Dasar Teori	10
2.2.1 Sistem Rekomendasi	10
2.2.2 <i>Content-Based Filtering</i>	12

2.2.3	Natural Language Processing (NLP)	15
2.2.4	<i>Data Preprocessing</i>	16
2.2.5	Semantic Textual Similarity (STS)	24
2.2.6	<i>Cosine similarity</i>	25
2.2.7	Transformer	26
2.2.8	Bidirectional Encoder Representations from Transformers	27
2.2.9	<i>Language-agnostic</i> BERT Sentence Embedding (LaBSE)	29
2.2.10	N-Gram	33
2.2.11	KeyBERT	34
2.2.12	Matriks Evaluasi	35
2.3	Jurnal Ilmiah	38
2.4	SINTA dan GARUDA	39
2.5	Integrasi Keilmuan	40
BAB III METODOLOGI PENELITIAN		44
3.1	Alur Penelitian	44
3.1.1	Studi Literatur	45
3.1.2	Data Collecting	46
3.1.3	Preprocessing Data	49
3.1.4	Ekstrasi Keyword dengan KeyBERT	62
3.1.5	<i>Data Splitting</i>	64
3.1.6	<i>Content-Based Filtering</i>	66
3.1.7	Analisa Hasil	74
BAB IV HASIL DAN PEMBAHASAN		76
4.1	Data Collecting	76
4.2	Preprocessing Data	80
4.3	Ekstrasi Keyword dengan KeyBERT	93

4.4	Data Splitting.....	97
4.5	Content Based Filtering.....	98
4.6	Analisa Hasil	103
BAB V Kesimpulan & Saran		111
5.1	Kesimpulan.....	111
5.2	Saran.....	112
Daftar Pustaka		113



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

Tabel 2.1 Daftar Penelitian Terdahulu.....	7
Tabel 2.2 Ilustrasi n-gram	33
Tabel 3.1 Sampel Data Jurnal dari Web SINTA.....	47
Tabel 3.2 Sampel Data Artikel dari Web Garuda	47
Tabel 3.3 Sampel Data Referensi dari API <i>Semantic Scholar</i>	48
Tabel 3.4 Contoh sampel penerapan <i>case folding</i>	52
Tabel 3.5 Contoh sampel penghapusan token numerik yang berdiri sendiri	53
Tabel 3.6 Contoh Penerapan Penghapusan Header pada Kolom <i>Abstrak</i>	54
Tabel 3.7 Contoh Penanganan Abstrak Ganda.....	55
Tabel 3.8 Contoh Penanganan Keyword Eksplisit pada Abstrak.....	55
Tabel 3.9 Contoh Sampel Setelah Deteksi Bahasa menggunakan <i>fastText</i>	60
Tabel 3.10 Contoh Sampel setelah Proses <i>Stopword Removal</i> Menggunakan Sastrawi untuk Bahasa Indonesia (id).....	61
Tabel 3.11 Contoh Sampel setelah Proses <i>Stopword Removal</i> menggunakan NLTK untuk Bahasa Inggris (en)	62
Tabel 3.12 Konfigurasi Parameter KeyBERT	62
Tabel 3.13 Sampel Dataset setelah Penambahan <i>Keyword</i>	64
Tabel 3.14 Skenario Kombinasi Bagian Artikel.....	69
Tabel 3.15 Skenario Pengujian Berdasarkan Skema Penelitian dan Model Embedding	71
Tabel 4.1 Daftar Jurnal yang Tidak Memiliki URL Garuda.....	77
Tabel 4.2 Daftar Jurnal Tanpa Publikasi Tahun 2022–2024.....	77
Tabel 4.3 Distribusi Jumlah Artikel pada Jurnal Teratas.....	78
Tabel 4.4 Perubahan Jumlah Artikel dan Jurnal Selama Tahap Preprocessing	80
Tabel 4.5 Statistik Distribusi Artikel per Jurnal	80
Tabel 4.6 Statistik Panjang Karakter Judul dan Abstrak	81
Tabel 4.7 Distribusi Missing Value pada Setiap Atribut Data.....	83
Tabel 4.8 Contoh Hasil Case Folding pada Kolom Judul	83
Tabel 4.9 Contoh Hasil Case Folding pada Kolom Abstrak	83
Tabel 4.10 Contoh Hasil Case Folding pada Kolom Referensess	84
Tabel 4.11 Hasil Penghapusan Token Numerik pada Teks	85

Tabel 4.12 Hasil Penghapusan Header Abstrak	85
Tabel 4.13 Hasil Penanganan Abstrak Ganda	86
Tabel 4.14 Hasil Penanganan Keyword Eksplisit pada Abstrak	86
Tabel 4.15 Ringkasan Data Sebelum dan Sesudah Duplicate Removal	87
Tabel 4.16 Contoh Data Duplikasi Berdasarkan Judul Artikel	87
Tabel 4.17 Contoh Data Duplikasi Berdasarkan Abstrak Artikel	88
Tabel 4.18 Data Terhapus Berdasarkan Batas Minimal Teks.....	89
Tabel 4.19 Daftar Jurnal dengan Jumlah Artikel di Bawah Ambang Batas	89
Tabel 4.20 Contoh Hasil Deteksi Bahasa pada Kolom Judul.....	90
Tabel 4.21 Contoh Hasil Deteksi Bahasa pada Kolom Abstrak.....	90
Tabel 4.22 Distribusi Bahasa pada Kolom Judul dan Abstrak	91
Tabel 4.23 Contoh Hasil Tokenization pada Teks Judul	92
Tabel 4.24 Daftar Custom Stopword.....	92
Tabel 4.25 Contoh Hasil Stopword Removal pada Teks Judul	92
Tabel 4.26 Contoh Hasil Ekstraksi Keyword Menggunakan KeyBERT	96
Tabel 4.27 Contoh Dataset Final	96
Tabel 4.28 Ringkasan Pembagian Data Latih dan Data Uji.....	97
Tabel 4.29 Distribusi Data Latih dan Data Uji pada Setiap Jurnal	97
Tabel 4.30 Hasil Evaluasi Kombinasi Metadata	98
Tabel 4.31 Perbandingan Performa Metode dan Model Embedding menggunakan Dataset GARUDA.....	101
Tabel 4.32 Input untuk Simulasi Perhitungan Cosine Similarity pada Skenario S1.3	108
Tabel 4.33 Hasil Simulasi Perhitungan Cosine Similarity pada Skenario S1.3 ..	109

DAFTAR GAMBAR

Gambar 2.1 Proses sistem rekomendasi jurnal berbasis CBF.....	12
Gambar 2.2 Visualisasi <i>bibliographic coupling</i>	13
Gambar 2.3 Arsitektur Transformer	26
Gambar 2.4 Arsitektur Model BERT	28
Gambar 2.5 Arsitektur LaBSE dengan dual-encoder	29
Gambar 2.6 Cara Kerja KeyBERT	34
Gambar 3.1 Alur Penelitian.....	45
Gambar 3.2 Alur Tahap Preprocessing.....	49
Gambar 3.3 Ilustrasi Alur Rekomendasi berdasarkan <i>Content-Based Filtering</i> ...	66
Gambar 4.1 Distribusi Jumlah Artikel per Jurnal.....	81
Gambar 4.2 Distribusi Jurnal berdasarkan Skor SINTA	82
Gambar 4.3 Visualisasi Frekuensi Unigram pada Dataset	93
Gambar 4.4 Visualisasi Frekuensi Bigram pada Dataset	94
Gambar 4.5 Visualisasi Distribusi Kata Kunci menggunakan Wordcloud	95
Gambar 4.6 Perbandingan Kinerja Skenario S1 pada Journal Accuracy	100
Gambar 4.7 Perbandingan Kinerja Skenario S1 pada MRR	100
Gambar 4.8 Perbandingan Kinerja Skenario S2 pada Journal Accuracy	102
Gambar 4.9 Perbandingan Kinerja Skenario S2 pada MRR	102

UIN SUNAN AMPEL
S U R A B A Y A

Daftar Pustaka

- Aditya, M. R., Ilham, M., Suriyanto, D. F., & Mappalotteng, A. M. (2025). Semantic Similarity Measurement Evaluation of KBBI Synonyms Using a Word Embedding Approach. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 14(2), 112–120.
- Andi, S. (2020). Pemanfaatan Science and Technology Index (SINTA) untuk Publikasi Karya Ilmiah dan Pencarian Jurnal Nasional Terakreditasi. *Media Pustakawan*, 27(1), 56–68.
- Anisatuzzumara. (2024). *Implementasi Latent Dirichlet Allocation (LDA) DAN K-Nearest Neighbors (KNN) pada Sistem Rekomendasi Jurnal Terindeks GARUDA*. Universitas Islam Sultan Agung.
- Avenberg, A. (2020). *Automatic language identification of short texts* (Issue September). UPPSALA UNIVERSITET. <http://www.teknat.uu.se/student>
- Chatfield, C. (1986). Exploratory data analysis. *European Journal of Operational Research*, 23(1), 5–13. https://doi.org/10.1007/978-3-319-39264-6_15
- Deshpande, A., Murahari, V., Kalyan, A., Chen, H., Graf, V., Chen, D., & Tech, G. (2023). *C-STs: Conditional Semantic Textual Similarity*. 5669–5690.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 1(Mlm), 4171–4186.
- Entrup, E., Eppelin, A., Ewerth, R., Hartwig, J., Tullney, M., Wohlgemuth, M., & Hoppe, A. (2024). Comparing different search methods for the open access journal recommendation tool B!SON. *International Journal on Digital Libraries*, 25(3), 505–516. <https://doi.org/10.1007/s00799-023-00372-3>
- Feng, F., Yang, Y., Cer, D., Arivazhagan, N., & Wang, W. (2022). Language-agnostic BERT Sentence Embedding. *Association for Computational Linguistics*, 1, 878–891. <https://doi.org/10.18653/v1/2022.acl-long.62>
- Gheewala, S., Xu, S., & Yeom, S. (2025). In-depth survey: deep learning in recommender systems—exploring prediction and ranking models, datasets, feature analysis, and emerging trends. In *Neural Computing and Applications*

- (Vol. 37, Issue 17). Springer London. <https://doi.org/10.1007/s00521-024-10866-z>
- Grootendorst, M. (2025). *KeyBERT Guides*. <https://maartengr.github.io/KeyBERT/guides/quickstart.html#prepare-embeddings>
- Gündoğan, E., Kaya, M., & Daud, A. (2023). Deep learning for journal recommendation system of research papers. *Scientometrics*, 128(1), 461–481. <https://doi.org/10.1007/s11192-022-04535-y>
- Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2017). Bag of Tricks for Efficient Text Classification. In M. Lapata, P. Blunsom, & A. Koller (Eds.), *Proceedings of the 15th Conference of the {E}uropean Chapter of the Association for Computational Linguistics: Volume 2, Short Papers* (pp. 427–431). Association for Computational Linguistics. <https://aclanthology.org/E17-2068/>
- Jurafsky, D., & Martin, J. H. (2025). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, with Language Models* (3rd (ed.)). stanford Edu.
- Kafi, A. Al, Banshal, S. K., Sultana, N., & Gupta, V. (2024). Source Recommendation System Using Context-based Classification: Empirical Study on Multi-level Ensemble Methods. *Journal of Scientometric Research*, 13(2), 475–484. <https://doi.org/10.5530/jsci.res.13.2.38>
- Khairani, U., Mutiawani, V., & Ahmadian, H. (2024). Pengaruh Tahapan Preprocessing Terhadap Model Indobert Dan Indobertweet Untuk Mendeteksi Emosi Pada Komentar Akun Berita Instagram. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(4), 887–894. <https://doi.org/10.25126/jtiik.1148315>
- Khurana, D., Koli, A., & Khatter, K. (2023). *Natural language processing : state of the art , current trends and challenges*. 3713–3744.
- Mulyani, E., Muhamad, F. P. B., & Cahyanto, K. A. (2021). Pengaruh N-Gram terhadap Klasifikasi Buku menggunakan Ekstraksi dan Seleksi Fitur pada Multinomial Naïve Bayes. *Jurnal Media Informatika Budidarma*, 5(1), 264. <https://doi.org/10.30865/mib.v5i1.2672>

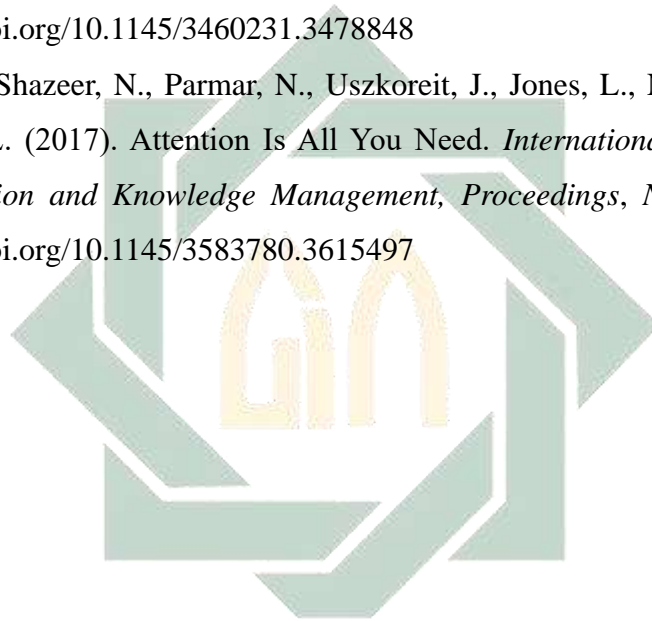
- Muriana, A. P., Sujainia, H., & Pratiwia, H. S. (2023). Sistem Rekomendasi Artikel sebagai Acuan Studi Literatur Menggunakan Metode N-Gram. *JURISTI (Jurnal Riset Sains Dan Teknologi Informatika)*, 01(1), 69–84. <https://doi.org/10.26418/juristi.v1i1.61170>
- Nigam, A., Dhruv, A., & Josephin S, F. J. (2021). Text Pre-Processing and Feature Extraction Using Nlp. *Www.Irjmets.Com @International Research Journal of Modernization in Engineering*, 03(06), 1550–1554. www.irjmets.com
- Noviandari, H. (2024). Optimalisasi Kinerja Publikasi Institusi Melalui Pengelolaan Jurnal Ilmiah Bereputasi. *Lentera Salam Institute*, 1(1), 32–43.
- Paaß, G., & Giesselbach, S. (2023). *Foundation Models for Natural Language Processing*.
- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using siamese BERT-networks. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, 3982–3992. <https://doi.org/10.18653/v1/d19-1410>
- Rianti, A., Majid, N. W. A., & Fauzi, A. (2024). Machine Learning Journal Article Recommendation System Using Content Based Filtering. *JUTI: Jurnal Ilmiah Teknologi Informasi*, 1–10. <https://doi.org/10.12962/j24068535.v22i1.a1193>
- Ricci, F., Rokach, L., & Shapira, B. (2010). *Recommender Systems Handbook* (Issue October). Springer, Boston, MA. <https://doi.org/10.1007/978-0-387-85820-3>
- Sabilillah, F. T., Winarno, S., & Abiyyi, R. B. (2024). Implementasi BERT dan Cosine Similarity untuk Rekomendasi Dosen Pembimbing berdasarkan Judul Tugas Akhir. *Edumatic: Jurnal Pendidikan Informatika*, 8(2), 585–594. <https://doi.org/10.29408/edumatic.v8i2.27791>
- Sadiartha, A. A. N. G. (2020). Best Practice: Penelitian Kualitatif dan Publikasi Ilmiah. In *Cakrawala Satria Mandiri*.
- Sanjaya, A., Setiawan, A. B., Mahdiyah, U., Farida, I. N., & Prasetyo, A. R. (2023). *PENGUKURAN KEMIRIPAN MAKNA MENGGUNAKAN COSINE SIMILARITY MEASUREMENT OF MEANING SIMILARITY USING*

COSINE SIMILARITY AND. 10(4), 747–752.
<https://doi.org/10.25126/jtiik.2023106864>

Soekamto, Y. S., Maryati, I., Christian, & Kurniawan, E. (2024). Improving Publishing: Extracting Keywords and Clustering Topics. *Journal of Applied Data Sciences*, 5(2), 412–423. <https://doi.org/10.47738/jads.v5i2.199>

Tamm, Y. (2021). Quality Metrics in Recommender Systems : Do We Calculate Metrics Consistently ? In *Proceedings of ACM Conference (Conference'17)* (Vol. 1, Issue 1). Association for Computing Machinery.
<https://doi.org/10.1145/3460231.3478848>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., N. Gomez, A., & Kaiser, Ł. (2017). Attention Is All You Need. *International Conference on Information and Knowledge Management, Proceedings, Nips*, 4752–4758.
<https://doi.org/10.1145/3583780.3615497>



UIN SUNAN AMPEL
S U R A B A Y A