SISTEM REKOMENDASI MITRA KOLABORASI PENELITIAN DOSEN BERDASARKAN KEMIRIPAN JUDUL PUBLIKASI MENGGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY

TUGAS AKHIR
SKRIPSI

SITI NURAINISA
41522010189

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026

lib.mercubuana.ac.id

SISTEM REKOMENDASI MITRA KOLABORASI PENELITIAN DOSEN BERDASARKAN KEMIRIPAN JUDUL PUBLIKASI MENGGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY

TUGAS AKHIR
SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana

SITI NURAINISA
41522010189

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2026

lib.mercubuana.ac.id

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama	:	Siti Nurainisa
NIM	:	41522010189
Fakultas/Program Studi	:	Fakultas Ilmu Komputer/Teknik Informatika

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:
"Sistem Rekomendasi Mitra Kolaborasi Penelitian Dosen Berdasarkan Kemiripan Judul Publikasi Menggunakan Sentence-BERT dan Cosine Similarity" adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

UNIVERSITAS
MERCU BUANA

Jakarta, 10 Agustus 2026



Siti Nurainisa

lib.mercubuana.ac.id

UNIVERSITAS
MERCU BUANA
AKREDITASI UNGGUL

PERNYATAAN SIMILARITY CHECK

SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh

The undersigned, hereby declare that the scientific work written by

Nama *Name*	: Siti Nurainisa
NIM *Student ID Number*	: 41522010189
Program Studi *Study Program*	: Teknik Informatika

Dengan Judul Tugas Akhir

The title

"SISTEM REKOMENDASI MITRA KOLABORASI PENELITIAN DOSEN BERDASARKAN KEMIRIPAN JUDUL PUBLIKASI MENGGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY"

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

Similarity Check has been carried out with the Turnitin system on the date

29 Juli 2026

dengan nilai persentase sebesar :

with a percentage value of

5%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer** Universitas Mercu Buana.

declared to meet the standards in accordance with the applicable regulations at the Faculty of Computer Science, Mercu Buana University.

File hasil cek similarity turnitin:

Turnitin similarity report file

https://docs.google.com/uc?id=1OXWq15rV0xMq0JJZ1tiUvHHRcoJLEx-&export=download&authuser=0&confirm=t&uuid=...

Jakarta, 29 Juli 2026
Admin Turnitin Fasilkom UMB

Agung Prawoto

Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama	:	Siti Nurainisa
NIM	:	41522010189
Fakultas/Program Studi	:	Fakultas Ilmu Komputer / Teknik Informatika
Judul Tugas Akhir	:	Sistem Rekomendasi Mitra Kolaborasi Penelitian Dosen Berdasarkan Kemiripan Judul Publikasi Menggunakan Sentence-BERT dan Cosine Similarity

Telah berhasil dipertahankan pada sidang tanggal 10 Agustus 2026 dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:
Pembimbing

Dr. Afiyati, S.Si., M.T.
NIDN/NUPTK: 0316106908

UNIVERSITAS
MERCU BUANA

Jakarta, 10 Agustus 2026
Mengetahui,

Dekan Fakultas Ilmu Komputer	Ketua Program Studi Teknik Informatika
Dr. Bambang Jokonowo, S.Si., MTI	Dr. Hadi Santoso, S.Kom., M.Kom
NIDN/NUPTK: 0320037002	NIDN/NUPTK: 0225067701

lib.mercubuana.ac.id

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Tugas Akhir ini. Penulisan Tugas Akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Teknik Informatika pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Tugas Akhir ini, sangatlah sulit bagi saya untuk menyelesaikan Tugas Akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercu Buana.
4. Ibu Dr. Afiyati, S.Si., M.T. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Kedua orang tua tercinta serta kakak-kakak penulis yang senantiasa memberikan doa, kasih sayang, dukungan, dan semangat kepada penulis. Terima kasih atas segala perhatian dan pengorbanan yang telah diberikan selama penulis menempuh pendidikan hingga menyelesaikan skripsi ini. Kehadiran dan dukungan keluarga menjadi salah satu kekuatan terbesar bagi penulis dalam menjalani setiap proses hingga sampai pada tahap ini.
6. Teman-teman penulis selama masa perkuliahan, khususnya Huda, Reni, dan Nadia, yang telah memberikan bantuan, dukungan, semangat, serta kebersamaan selama menjalani perkuliahan hingga proses penyusunan skripsi ini. Terima kasih telah menjadi bagian dari perjalanan penulis, baik dalam menghadapi berbagai kesulitan maupun berbagi cerita dan pengalaman selama masa kuliah.
7. Sahabat-sahabat penulis sejak masa sekolah, yaitu Rahma, Naya, dan Via yang telah menjadi teman sejak SMP, serta Sabrina, dan Dela yang telah menjadi teman sejak SMA. Terima kasih atas dukungan, perhatian,

lib.mercubuana.ac.id

kebersamaan, serta semangat yang diberikan kepada penulis selama ini. Meskipun memiliki kesibukan masing-masing, persahabatan yang tetap terjalin hingga saat ini menjadi hal yang berarti bagi penulis.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tugas Akhir ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 10 Agustus 2026

Siti Nurainisa

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR DI REPOSITORI UMB

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama	:	Siti Nurainisa
NIM	:	41522010189
Fakultas/Program Studi	:	Fakultas Ilmu Komputer / Teknik Informatika
Judul Laporan Skripsi	:	Sistem Rekomendasi Mitra Kolaborasi Penelitian Dosen Berdasarkan Kemiripan Judul Publikasi Menggunakan Sentence-BERT dan Cosine Similarity

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif** (***Non-exclusive Royalty-Free Right***) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).
Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 10 Agustus 2026

Yang menyatakan,



Siti Nurainisa

lib.mercubuana.ac.id

SISTEM REKOMENDASI MITRA KOLABORASI PENELITIAN DOSEN BERDASARKAN KEMIRIPAN JUDUL PUBLIKASI MENGGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY

SITI NURAINISA

ABSTRAK

Penelitian ini membangun sistem rekomendasi mitra kolaborasi penelitian dosen berdasarkan kemiripan judul publikasi menggunakan Sentence-BERT dan cosine similarity, dengan TF-IDF sebagai metode pembanding. Data diperoleh melalui *crawling* Google Scholar dari 75 dosen pada Fakultas Ilmu Komputer di satu universitas. Hasil pengumpulan data menghasilkan 2.625 rekaman mentah. Setelah deduplikasi awal, validasi tahun, penyatuan identitas penulis, normalisasi judul, dan konsolidasi lintas profil, diperoleh 2.226 publikasi unik, 2.889 relasi kepenulisan, dan 75 dosen yang terhubung dengan publikasi. Evaluasi dilakukan menggunakan pemisahan temporal, yaitu publikasi sampai tahun 2023 sebagai data historis dan publikasi tahun 2024–2025 sebagai data pengujian. Setiap judul dikodekan secara terpisah menggunakan paraphrase-multilingual-MiniLM-L12-v2 menjadi vektor 384 dimensi. Skor kandidat dihitung berdasarkan rata-rata tiga bukti publikasi terbaik. Pada protokol seluruh kolaborator masa depan dengan K=5, Sentence-BERT memperoleh Precision@5 sebesar 0,1968, Recall@5 sebesar 0,1601, dan Hit Rate@5 sebesar 0,6290. TF-IDF memperoleh nilai masing-masing sebesar 0,2548, 0,2078, dan 0,7581. Pada evaluasi menggunakan judul baru, Sentence-BERT menghasilkan Recall@10 sebesar 0,4039 dan Hit Rate@10 sebesar 0,5000, sedangkan TF-IDF menghasilkan nilai masing-masing sebesar 0,3777 dan 0,4627. Hasil penelitian menunjukkan bahwa TF-IDF lebih sesuai untuk menghasilkan rekomendasi mitra kolaborasi pada dataset ini, sedangkan Sentence-BERT lebih baik pada beberapa metrik ketika masukan hanya berupa satu judul baru. Prototipe Streamlit digunakan untuk mendemonstrasikan rekomendasi berdasarkan profil dosen dan judul penelitian baru.

Kata kunci: sistem rekomendasi, mitra kolaborasi, Sentence-BERT, cosine similarity, TF-IDF, kepenulisan bersama.

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

RECOMMENDATION SYSTEM FOR LECTURER RESEARCH COLLABORATION PARTNERS BASED ON PUBLICATION TITLE SIMILARITY USING SENTENCE-BERT AND COSINE SIMILARITY

SITI NURAINISA

ABSTRACT

This study develops a lecturer research collaboration partner recommendation system based on publication title similarity using Sentence-BERT and cosine similarity, with TF-IDF as a comparison method. Data were collected through Google Scholar crawling from 75 lecturers affiliated with the Faculty of Computer Science at a single university. The data collection process produced 2,625 raw records. After initial deduplication, year validation, author identity resolution, title normalization, and cross-profile consolidation, the final dataset consisted of 2,226 unique publications, 2,889 authorship relationships, and 75 lecturers linked to the publications. The evaluation was conducted using a temporal split, with publications up to 2023 used as historical data and publications from 2024–2025 used as test data. Each title was encoded separately using the paraphrase-multilingual-MiniLM-L12-v2 model into a 384-dimensional vector. Candidate scores were calculated based on the average of the three highest publication-level similarity scores. Under the all-future-collaborators protocol with K=5, Sentence-BERT achieved a Precision@5 of 0.1968, a Recall@5 of 0.1601, and a Hit Rate@5 of 0.6290. TF-IDF achieved corresponding scores of 0.2548, 0.2078, and 0.7581. In the evaluation using new titles, Sentence-BERT achieved a Recall@10 of 0.4039 and a Hit Rate@10 of 0.5000, whereas TF-IDF achieved 0.3777 and 0.4627, respectively. The results indicate that TF-IDF is more suitable for generating recommendations that align with future co-authorship patterns in this dataset, whereas Sentence-BERT performs better on several metrics when the input consists of only a single new title. A Streamlit prototype was used to demonstrate recommendations based on lecturer profiles and new research titles.

Keywords: *recommender system, collaboration partner, Sentence-BERT, cosine similarity, TF-IDF, co-authors*

lib.mercubuana.ac.id

DAFTAR ISI

HALAMAN SAMPUL **0**

HALAMAN JUDUL ..... **i**

HALAMAN PERNYATAAN KARYA SENDIRI ..... **ii**

HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN ..... **iii**

HALAMAN PENGESAHAN ..... **iv**

KATA PENGANTAR ..... **v**

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB ..... **vii**

ABSTRAK ..... **viii**

ABSTRACT ..... **ix**

DAFTAR ISI ..... **x**

DAFTAR TABEL ..... **xiii**

DAFTAR GAMBAR ..... **xiv**

DAFTAR LAMPIRAN ..... **xv**

BAB I **1**

PENDAHULUAN **1**

1. 1 Latar Belakang 1

1. 2 Perumusan Masalah 2

1.3 Tujuan Penelitian 3

1.4 Manfaat Penelitian 3

1.5 Batasan Masalah 3

BAB II **5**

TINJAUAN PUSTAKA **5**

2. 1 Penelitian Terdahulu 5

2. 2 Teori Pendukung 17

2.2.1 Sistem Rekomendasi 17

2.2.2 *Expert finding* dan Rekomendasi Mitra 18

2.2.3 Publikasi dan *Co-authorship* 18

2.2.4 Judul Publikasi sebagai Representasi Topik 19

lib.mercubuana.ac.id

2.2.5 Pembersihan dan Normalisasi Teks 19
2.2.6 TF-IDF 20
2.2.7 Sentence-BERT 20
2.2.8 Cosine similarity 21
2.2.9 Evaluasi Top-K 22
2.2.10 *Web crawling* Google Scholar 23
2.2.11 Streamlit 24
BAB III 25
METODE PENELITIAN 25
3. 1 Jenis Penelitian 25
3. 2 Tahapan Penelitian 25
3. 3 Pengumpulan Data Google Scholar 26
3. 4 Penggabungan dan Pemeriksaan Data Awal 30
3. 5 Penyatuan Identitas Dosen dan Penulis 30
3. 6 Validasi Tahun, Normalisasi Judul, dan Deduplikasi Final 30
3. 7 Pemisahan Data Berdasarkan Waktu 31
3. 8 Pembentukan Representasi Sentence-BERT 32
3. 9 Pembentukan Metode pembanding TF-IDF 32
3. 10 Perhitungan Kemiripan dan Agregasi Skor 33
3. 11 Pembentukan *Ground truth* 34
3. 12 Metrik Evaluasi 35
3. 13 Konfigurasi Eksperimen 35
3.14 Implementasi Prototipe Streamlit 36
BAB IV 38
PEMBAHASAN 38
4. 1 Hasil Pengambilan Data Google Scholar 38
4. 2 Penggabungan dan Pemeriksaan Awal 38

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

4. 3 Hasil Penyatuan Identitas Penulis 40
4. 4 Validasi Tahun dan Konsolidasi Publikasi 41
4. 5 Analisis Deskriptif Dataset Bersih 43
4. 6 Hasil Pemisahan Temporal 45
4. 7 Hasil Pembentukan Representasi 46
4. 8 Contoh Hasil Rekomendasi 47
4. 9 *Ground Truth* Kolaborasi Temporal 49
4. 10 Evaluasi Seluruh Kolaborator Masa Depan 49
4. 11 Evaluasi Kolaborator Baru 52
4. 12 Evaluasi Berdasarkan Judul Baru 54
4. 13 Pembahasan Perbandingan Metode 55
4. 14 Implementasi Prototipe Streamlit 58
BAB V 62
KESIMPULAN DAN SARAN 62
5.1 Kesimpulan 62
5.2 Saran 63
DAFTAR PUSTAKA 64
LAMPIRAN 67

UNIVERSITAS MERCU BUANA

lib.mercubuana.ac.id

DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu 5
Tabel 3. 1 Atribut data hasil *crawling* 28
Tabel 3. 2 Konfigurasi pengambilan data Google Scholar 29
Tabel 3. 3 Rancangan pemisahan temporal 31
Tabel 3. 4 Konfigurasi model dan evaluasi 35
Tabel 4. 1 Ringkasan data sebelum pembersihan 39
Tabel 4. 2 Hasil audit tahun publikasi 42
Tabel 4. 3 Ringkasan hasil pembersihan dan konsolidasi 42
Tabel 4. 4 Komposisi pemisahan temporal 45
Tabel 4. 5 Representasi yang digunakan 47
Tabel 4. 6 Contoh rekomendasi berdasarkan profil dosen 47
Tabel 4. 7 Contoh rekomendasi berdasarkan judul baru 48
Tabel 4. 8 Ringkasan *ground truth* temporal 49
Tabel 4. 9 Hasil evaluasi seluruh kolaborator masa depan 50
Tabel 4. 10 Hasil evaluasi kolaborator baru 52
Tabel 4. 11 Hasil evaluasi judul baru 54
Tabel 4. 12 Perbandingan waktu komputasi Sentence-BERT dan TF-IDF 57
Tabel 4. 13 Pengujian fungsional prototipe Streamlit 60

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR GAMBAR

Gambar 3. 1 Tahapan penelitian 26
Gambar 3. 2 Alur pengambilan data Google Scholar 28
Gambar 3. 3 Rancangan pemodelan dan evaluasi..... 34
Gambar 4. 1 Dampak deduplikasi awal 40
Gambar 4. 2 Distribusi status pemeriksaan pemetaan penulis..... 41
Gambar 4. 3 Status validasi tahun publikasi 42
Gambar 4. 4 Distribusi jumlah publikasi berdasarkan tahun 44
Gambar 4. 5 Tiga puluh dosen dengan jumlah publikasi terbanyak..... 44
Gambar 4. 6 Jejaring kolaborasi dosen 45
Gambar 4. 7 Komposisi pemisahan temporal 46
Gambar 4. 8 Perbandingan Precision@K pada seluruh kolaborator masa depan. 51
Gambar 4. 9 Perbandingan Recall@K pada seluruh kolaborator masa depan..... 51
Gambar 4. 10 Perbandingan Hit Rate@K pada seluruh kolaborator masa depan 52
Gambar 4. 11 Perbandingan Hit Rate@K pada kolaborator baru..... 53
Gambar 4. 12 Perbandingan Hit Rate@K pada evaluasi judul baru..... 55
Gambar 4. 13 Tampilan rekomendasi berdasarkan profil dosen 58
Gambar 4. 14 Hasil rekomendasi berdasarkan profil dosen 59
Gambar 4. 15 Tampilan rekomendasi berdasarkan judul baru 59
Gambar 4. 16 Hasil rekomendasi berdasarkan judul baru 60

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id

DAFTAR LAMPIRAN

Lampiran 1. Kartu Asistensi 67

Lampiran 2. Curriculum Vitae 68

Lampiran 3. Surat Pernyataan & Pengalihan HAKI 69

Lampiran 4. Sertifikat BNSP 71

UNIVERSITAS
MERCU BUANA

lib.mercubuana.ac.id