

**SISTEM REKOMENDASI KANDIDAT PENCARI KERJA
MENGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

EGA SAPUTRA HUTAGALUNG

21.11.4345

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**SISTEM REKOMENDASI KANDIDAT PENCARI KERJA
MENGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY**

SKRIPSI

Untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Informatika



disusun oleh

EGA SAPUTRA HUTAGALUNG

21.11.4345

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**SISTEM REKOMENDASI KANDIDAT PENCARI KERJA
MENGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY**

yang disusun dan diajukan oleh

Ega Saputra Hutagalung

21.11.4345

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 11 Agustus 2025

Dosen Pembimbing,


Mulia Sulistiyono, S.Kom., M.Kom.

NIK. 190302248

HALAMAN PENGESAHAN

SKRIPSI

**SISTEM REKOMENDASI KANDIDAT PENCARI KERJA
MENGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY**

yang disusun dan diajukan oleh

Ega Saputra Hutagalung

21.11.4345

Telah dipertahankan di depan Dewan Penguji
pada tanggal 11 Agustus 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Tri Susanto, M.Kom.
NIK. 190302115

Dina Maulina, S.Kom., M.Kom.
NIK. 190302250

Mulia Sulistiyono, S.Kom., M.Kom.
NIK. 190302248

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 11 Agustus 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ega Saputra Hutagalung
NIM : 21.11.4345

Menyatakan bahwa Skripsi dengan judul berikut:

SISTEM REKOMENDASI KANDIDAT PENCARI KERJA MENGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY

Dosen Pembimbing: Mulia Sulistiyono, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta,

Yang Menyatakan.



Ega Saputra Hutagalung

HALAMAN PERSEMBAHAN

Dengan mengucapkan syukur ke hadirat Allah SWT atas segala rahmat, karunia, dan kemudahan-Nya, penulis dapat menyelesaikan tugas akhir ini. Karya ini tidak lepas dari peran dan dukungan banyak pihak yang telah memberikan semangat, motivasi, dan doa yang tulus selama proses studi hingga penyusunan skripsi ini. Oleh karena itu, dengan penuh hormat dan rasa syukur, karya ini penulis persembahkan kepada:

1. Tuhan, yang senantiasa menjadi tempat berkeluh kesah atas proses berjalannya dalam penyelesaian studi ini.
2. Kedua orang tua yang selalu memberikan motivasi, semangat, cinta, serta pengorbanan yang menjadi sumber semangat dalam penyelesaian studi ini.
3. Sinta Juliyanti, yang selalu hadir sebagai sumber semangat, ketenangan, dan kekuatan di setiap langkah perjuangan. Terima kasih telah sabar, mendukung, dan percaya terhadap proses penyelesaian studi ini.
4. Bapak Mulia Sulistiyono, S.Kom., M.Kom., selaku dosen pembimbing, atas segala ilmu, arahan, dan bimbingan yang sangat berarti.
5. Seluruh dosen dan staf Program Studi Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta, yang telah memberikan bekal ilmu pengetahuan, pengalaman, dan nilai-nilai pembelajaran selama masa studi.
6. Rekan-rekan seperjuangan, yang selalu memberikan rasa kebersamaan, kerja sama, dan saling dukung-mendukung selama proses penyelesaian studi ini.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “SISTEM REKOMENDASI KANDIDAT PENCARI KERJA MENGGUNAKAN SENTENCE-BERT DAN COSINE SIMILARITY” ini dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta.

Ucapan terima kasih penulis disampaikan kepada:

1. Prof. Dr. M. Suyanto., M.M., selaku Rektor Universitas Amikom Yogyakarta.
2. Prof. Dr. Kusrini, M.Kom., selaku Dekan Fakultas Ilmu Komputer.
3. Eli Pujastuti, M.Kom., selaku Ketua Program Studi Informatika
4. Bapak Mulia Sulistiyono, S.Kom., M.Kom., selaku dosen pembimbing, atas segala ilmu, arahan, dan bimbingan yang sangat berarti.
5. Seluruh dosen, staf pengajar, dan tenaga kependidikan di Program Studi Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta, yang telah memberikan ilmu dan pengalaman berharga selama masa studi.
6. Kedua orang tua dan keluarga tercinta, yang senantiasa mendoakan dan memberikan dukungan moril maupun material tanpa henti.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, baik dari segi isi maupun penyajiannya. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan dapat menjadi referensi bagi penelitian selanjutnya.

Yogyakarta, 7 Agustus 2025

Penulis

DAFTAR ISI

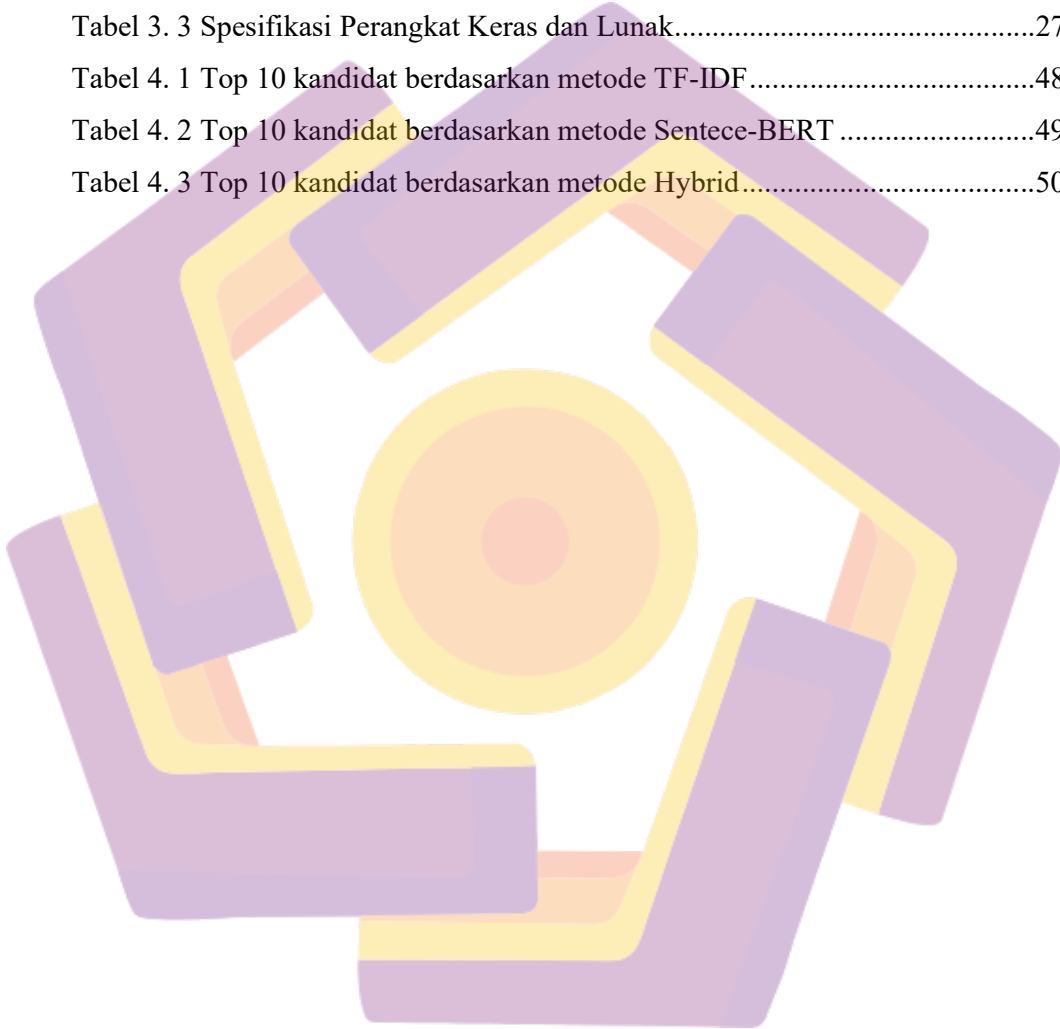
HALAMAN JUDUL	i
HALAMAN JUDUL	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
INTISARI	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	11

2.2.1	Sistem Rekomendasi	11
2.2.2	Content-Based Filtering	11
2.2.3	Preprocessing	13
2.2.4	TF-IDF	14
2.2.5	Sentence-BERT.....	15
2.2.6	Cosine Similarity	16
2.2.7	Streamlit.....	17
BAB III METODE PENELITIAN		19
3.1	Objek Penelitian.....	19
3.2	Alur Penelitian	19
3.2.1	Persiapan Dataset	21
3.2.2	Preprocessing	21
3.2.3	Ekstraksi Fitur.....	22
3.2.4	Perhitungan Cosine Similarity	22
3.2.5	Evaluasi Cosine Similarity.....	23
3.3	Alat dan Bahan.....	24
3.3.1	Data penelitian	24
3.3.2	Alat/Instrumen	27
3.3.3	Implementasi Sistem.....	27
BAB IV HASIL DAN PEMBAHASAN		29
4.1	Persiapan Dataset	29
4.1.1	Pengumpulan Data	29
4.1.2	Pemilihan Data.....	29
4.2	Persiapan Sistem	29
4.2.1	Pemanggilan Pustaka	29

4.2.2	Inisialisasi Antarmuka Streamlit.....	31
4.2.3	Upload Dataset.....	31
4.3	Preprocessing Data.....	33
4.3.1	Pembersihan Kolom Teks	33
4.3.2	Penggabungan Kolom untuk Representasi Semantik	36
4.3.3	Visualisasi Dampak Preprocessing	37
4.4	Hasil Ekstraksi Fitur	38
4.4.1	Ekstraksi dengan TF-IDF	39
4.4.2	Ekstraksi dengan Sentence-BERT	40
4.4.3	Skor Hybrid (Gabungan TF-IDF dan SBERT).....	41
4.5	Evaluasi Cosine Similarity dan Rekomendasi Kandidat.....	42
4.6	Perbandingan Metode TF-IDF, Sentence-BERT, dan Hybrid.....	45
4.7	Kelemahan Sistem	52
BAB V PENUTUP		53
5.1	Kesimpulan	53
5.2	Saran	54
REFERENSI		55

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	8
Tabel 3. 1 Contoh Struktur Dataset Kandidat Pencari Kerja	25
Tabel 3. 2 Struktur Dataset Lowongan Pekerjaan.....	26
Tabel 3. 3 Spesifikasi Perangkat Keras dan Lunak.....	27
Tabel 4. 1 Top 10 kandidat berdasarkan metode TF-IDF.....	48
Tabel 4. 2 Top 10 kandidat berdasarkan metode Sentece-BERT	49
Tabel 4. 3 Top 10 kandidat berdasarkan metode Hybrid.....	50

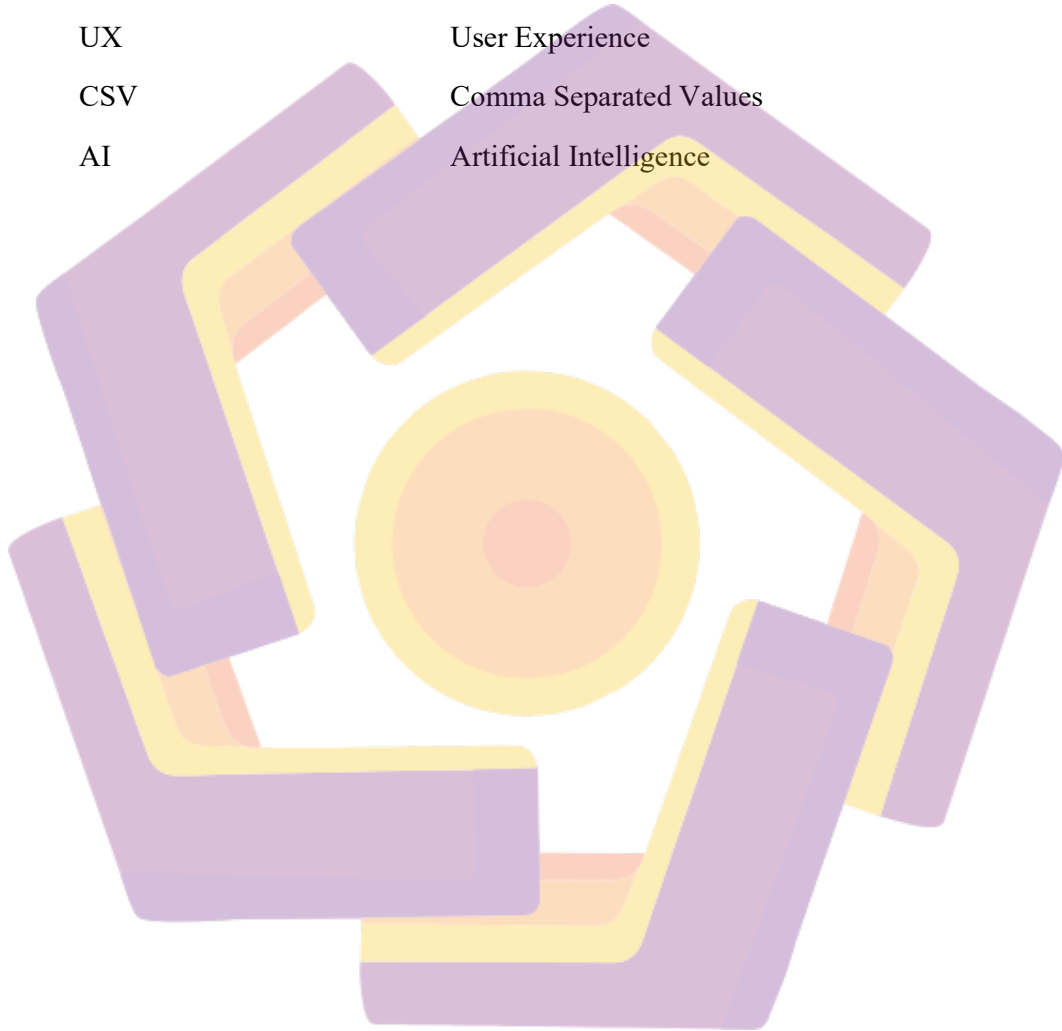


DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi konsep content-based filtering	12
Gambar 3. 1 Diagram Alur Penelitian	20
Gambar 4. 1 Sebelum Upload Dataset	33
Gambar 4. 2 Sesudah Upload Dataset.....	33
Gambar 4. 3 Pembersihan dan Preprocessing pada kedua dataset.....	35
Gambar 4. 4 Wordcloud Sebelum Preprocessing Pada Dataset Kandidat.....	38
Gambar 4. 5 Wordcloud Setelah Preprocessing Pada Dataset Kandidat	38
Gambar 4. 6 Wordcloud Sebelum Preprocessing Pada Dataset Lowongan	38
Gambar 4. 7 Wordcloud Setelah Preprocessing Pada Dataset Lowongan.....	38
Gambar 4. 8 Visualisasi Skor Kandidat	44
Gambar 4. 9 Heatmap Kemiripan Antar Kandidat	44
Gambar 4. 10 Hasil grafik berdasarkan metode TF-IDF.	45
Gambar 4. 11 Hasil grafik berdasarkan metode Sentence-BERT.....	46
Gambar 4. 12 Hasil grafik berdasarkan metode hybrid.	47

DAFTAR LAMBANG DAN SINGKATAN

TF-IDF	Term Frequency - Inverse Document Frequency
SBERT	Sentence-BERT
UI	User Interface
UX	User Experience
CSV	Comma Separated Values
AI	Artificial Intelligence



INTISARI

Proses pencarian kandidat yang tepat dalam rekrutmen kerja sering kali memakan waktu dan tidak efisien apabila dilakukan secara manual. Untuk mengatasi permasalahan tersebut, penelitian ini mengembangkan sistem rekomendasi kandidat pencari kerja dengan memanfaatkan metode Sentence-BERT dan Cosine Similarity. Dataset yang digunakan terdiri dari profil kandidat serta deskripsi lowongan pekerjaan. Setelah dilakukan proses preprocessing, data diekstraksi menggunakan Sentence-BERT untuk menghasilkan representasi vektor. Kemudian, tingkat kemiripan antara kandidat dan pekerjaan dihitung menggunakan Cosine Similarity. Sistem ini diimplementasikan melalui antarmuka web interaktif menggunakan Streamlit, sehingga pengguna dapat mengunggah dataset dan melihat rekomendasi kandidat terbaik berdasarkan kecocokan skor. Hasil evaluasi menunjukkan bahwa metode ini mampu memberikan rekomendasi kandidat secara kontekstual dan relevan terhadap posisi yang dicari. Dengan demikian, sistem ini dapat membantu proses seleksi kandidat menjadi lebih efisien dan objektif.

Kata kunci: Sistem Rekomendasi, Sentence-BERT, Cosine Similarity, Kandidat Pekerjaan, Streamlit

ABSTRACT

The process of finding suitable candidates in recruitment is often time-consuming and inefficient when conducted manually. To address this issue, this study develops a job candidate recommendation system utilizing the Sentence-BERT method and Cosine Similarity. The dataset consists of candidate profiles and job descriptions. After preprocessing, Sentence-BERT is used to extract vector representations, and Cosine Similarity is applied to calculate the similarity between candidates and job requirements. The system is implemented through an interactive web interface using Streamlit, allowing users to upload datasets and view the top candidate recommendations based on similarity scores. Evaluation results show that this method is capable of providing contextual and relevant candidate matches for the desired positions. Therefore, this system can help streamline the candidate selection process and improve objectivity.

Keyword: Recommendation System, Sentence-BERT, Cosine Similarity, Job Candidates, Streamlit