

# Implementasi Content-Based Filtering Untuk Sistem Rekomendasi Pekerjaan Berbasis *Curriculum Vitae* (CV)

Winna Dhestyan Putri<sup>1</sup>, Dwi Arman Prasetya<sup>2\*</sup>, Amri Muhaimin<sup>3</sup>

<sup>1,2,3</sup>Sains Data, Universitas Pembangunan Nasional “Veteran” Jawa Timur

[122083010015@student.upnjatim.ac.id](mailto:122083010015@student.upnjatim.ac.id)

[amri.muhamin.stat@upnjatim.ac.id](mailto:amri.muhamin.stat@upnjatim.ac.id)

\*Corresponding author email: [arman.prasetya.sada@upnjatim.ac.id](mailto:arman.prasetya.sada@upnjatim.ac.id)

**Abstrak**— Proses pencarian kerja di era digital menghadapi tantangan berupa tingginya jumlah pelamar dan rendahnya tingkat kesesuaian antara profil pelamar dengan kebutuhan perusahaan. Permasalahan ini diperparah oleh kompleksitas data teks tidak terstruktur pada *Curriculum Vitae* (CV) dan deskripsi pekerjaan, sehingga proses pencocokan menjadi kurang optimal. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi pekerjaan berbasis *Content-Based Filtering* (CBF) dengan memanfaatkan representasi teks menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) dan pengukuran kemiripan menggunakan cosine similarity. Dataset yang digunakan terdiri dari dokumen CV dalam format PDF serta data lowongan pekerjaan yang dikumpulkan dari platform LinkedIn. Tahapan penelitian meliputi ekstraksi teks, *preprocessing* menggunakan teknik *Natural Language Processing* (NLP), pembentukan vektor TF-IDF, serta perhitungan skor kemiripan untuk menghasilkan rekomendasi pekerjaan dalam bentuk Top-N. Sistem diimplementasikan menggunakan Streamlit untuk menyediakan antarmuka yang interaktif dan mudah digunakan. Hasil evaluasi menunjukkan bahwa sistem mampu menghasilkan rekomendasi yang relevan dengan rata-rata skor relevansi sebesar 4,2 dari 5 dan tingkat kepuasan pengguna sebesar 4,5 dari 5, dengan 90,5% responden menyatakan puas. Dengan demikian, pendekatan CBF berbasis TF-IDF terbukti efektif dalam meningkatkan kualitas pencocokan pekerjaan secara praktis.

**Kata Kunci** — Sistem rekomendasi, *Natural Language Processing*, *Content-Based Filtering*, Cosine Similarity, *Curriculum Vitae*, LinkedIn, TF-IDF

## I. PENDAHULUAN

Proses pencarian kerja dan perekrutan karyawan telah menjadi tantangan utama di era digital. Meskipun platform daring memperluas akses terhadap lowongan pekerjaan, tingkat persaingan meningkat secara signifikan. Laporan *LinkedIn Work Change Report 2025* menunjukkan bahwa sekitar 58% tenaga kerja global berencana mencari pekerjaan baru, sementara 37% pencari kerja merasa proses pencarian semakin sulit akibat tingginya jumlah pelamar pada setiap lowongan [1]. Seiring meningkatnya persaingan di pasar kerja, perusahaan juga mulai mengubah strategi rekrutmen untuk memperoleh kandidat yang sesuai dengan kebutuhan industri yang terus berkembang. Penelitian terbaru menunjukkan bahwa pada bidang yang mengalami kekurangan talenta, seperti *Artificial Intelligence* (AI) dan *green jobs*, perusahaan mulai menerapkan pendekatan *skill-based hiring* dengan lebih menekankan

keterampilan yang dimiliki kandidat dibandingkan kualifikasi formal. Peningkatan permintaan terhadap pekerjaan berbasis AI diikuti dengan menurunnya persyaratan gelar pendidikan formal pada banyak lowongan pekerjaan, yang menunjukkan bahwa kompetensi dan keterampilan semakin menjadi faktor utama dalam proses rekrutmen tenaga kerja modern [2].

Meskipun demikian, berbagai penelitian menunjukkan bahwa ketidaksesuaian antara kualifikasi, pendidikan, dan keterampilan yang dimiliki pencari kerja dengan kebutuhan perusahaan masih menjadi permasalahan yang signifikan di pasar tenaga kerja. Fenomena *qualification mismatch* dapat menyebabkan ketidakseimbangan antara karakteristik pencari kerja dan kebutuhan lowongan yang tersedia sehingga memengaruhi efektivitas proses penempatan tenaga kerja [3]. Selain itu, *skill mismatch* diketahui berkontribusi terhadap meningkatnya tingkat pengangguran, tingginya perputaran tenaga kerja (*labour turnover*), serta rendahnya efisiensi pasar kerja [4].

Permasalahan tersebut juga diperkuat oleh adanya *field-of-study mismatch* dan *overqualification*, yaitu kondisi ketika pekerja bekerja pada bidang yang tidak sesuai dengan latar belakang pendidikan atau memiliki tingkat kualifikasi yang melebihi kebutuhan pekerjaan. Kondisi ini dapat berdampak pada kualitas pencocokan pekerjaan, produktivitas tenaga kerja, serta efisiensi pasar kerja secara keseluruhan [5]. Di sisi lain, analisis terhadap data lowongan pekerjaan menunjukkan bahwa karakteristik pendidikan dan keterampilan yang dipersyaratkan perusahaan sering kali tidak selaras dengan karakteristik tenaga kerja yang tersedia, sehingga memperbesar potensi terjadinya *mismatch* dalam proses rekrutmen [6]. Oleh karena itu, diperlukan mekanisme yang mampu membantu proses pencocokan antara profil pelamar dan kebutuhan perusahaan secara lebih efektif.

Salah satu pendekatan yang berkembang untuk mengatasi permasalahan tersebut adalah pemanfaatan *Job Recommender System*. Tinjauan literatur terbaru menunjukkan bahwa sistem rekomendasi pekerjaan memiliki potensi besar dalam membantu pencari kerja menemukan lowongan yang relevan sekaligus membantu perusahaan memperoleh kandidat yang sesuai dengan kebutuhan posisi yang ditawarkan. Namun demikian, tantangan terkait relevansi rekomendasi, pengolahan data tidak terstruktur, personalisasi hasil rekomendasi, serta

kualitas proses pencocokan masih menjadi isu yang perlu dikembangkan lebih lanjut [7].

Dalam praktik rekrutmen, proses pencocokan kandidat umumnya dilakukan dengan membandingkan informasi yang terdapat pada *Curriculum Vitae* (CV) pelamar dengan kualifikasi dan kebutuhan yang tercantum pada deskripsi lowongan pekerjaan. Oleh karena itu, kualitas representasi dan analisis terhadap kedua sumber informasi tersebut menjadi faktor penting dalam menghasilkan rekomendasi pekerjaan yang relevan. Permasalahan utama terletak pada kompleksitas data teks dalam CV dan deskripsi pekerjaan. Variasi istilah, perbedaan format penulisan, serta ambiguitas bahasa menyebabkan proses pencocokan menjadi kurang optimal [8]. Selain itu, proses transformasi data teks ke dalam bentuk terstruktur sering kali menyebabkan hilangnya informasi penting, sehingga menurunkan kualitas rekomendasi yang dihasilkan [9]. Kondisi ini menyebabkan sistem yang ada masih menghasilkan rekomendasi yang kurang relevan serta belum mampu menangani data secara efektif.

Untuk mengatasi permasalahan tersebut, sistem rekomendasi pekerjaan berbasis *Content-Based Filtering* (CBF) menjadi salah satu pendekatan yang relevan. CBF bekerja dengan menganalisis kesamaan konten antara profil pengguna dan item yang direkomendasikan tanpa bergantung pada data pengguna lain [10]. Dalam penelitian ini, representasi teks dilakukan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), yang mampu memberikan bobot lebih besar pada kata-kata penting dalam dokumen [9]. Selanjutnya, tingkat kesamaan antara CV dan deskripsi pekerjaan dihitung menggunakan *cosine similarity* untuk menghasilkan skor relevansi [11].

Meskipun pendekatan CBF telah banyak digunakan, sebagian besar penelitian sebelumnya masih berfokus pada evaluasi model tanpa integrasi ke dalam aplikasi yang dapat digunakan secara langsung oleh pengguna. Selain itu, pemanfaatan data CV dalam format tidak terstruktur seperti PDF serta penyajian hasil rekomendasi secara interaktif masih relatif terbatas. Oleh karena itu, penelitian ini mengusulkan penerapan metode *Content-Based Filtering* berbasis TF-IDF dan *cosine similarity* untuk menghasilkan rekomendasi pekerjaan yang relevan berdasarkan CV pelamar. Sistem yang dibangun mampu mengolah data CV dalam format PDF, melakukan *preprocessing* teks secara sistematis, serta menyajikan hasil rekomendasi dalam bentuk peringkat (*Top-N Recommendation*) melalui antarmuka berbasis Streamlit. Dengan demikian, penelitian ini diharapkan dapat meningkatkan efektivitas dan efisiensi proses pencarian kerja secara praktis dan aplikatif.

#### A. Penelitian Terdahulu

Tabel I. PENELITIAN TERDAHULU TENTANG SISTEM REKOMENDASI PEKERJAAN

Peneliti	Metode	Dataset	Hasil Utama
Ro'uf et al. (2023) [11]	<i>Content-Based Filtering</i> (CBF) + <i>MLP Classifier</i>	Data Alumni (1.120)	Menunjukkan performa sangat tinggi (akurasi hingga 100%), namun sensitif terhadap pembagian data; metode efektif dalam memodelkan kecocokan pekerjaan
Alya Fitria et.al (2024) [12]	<i>Content-Based Filtering</i> (CBF)	Lowongan kerja dari Jobstreet (437) dan profil pencari kerja dari LinkedIn (100)	CBF menghasilkan rekomendasi relevan dan cukup akurat dengan recall 0,6397 dan presisi 0,53, berdasarkan analisis atribut seperti lokasi, posisi, keahlian, dan deskripsi pekerjaan
Koloman et al. (2023) [13]	<i>Content-Based Filtering</i> (CBF)	Data pekerjaan bidang IT (LinkedIn dan Glassdoor)	CBF mampu memberikan rekomendasi yang relevan dan membantu pencari kerja menemukan peluang sesuai keterampilan mereka.

Penelitian ini menawarkan pendekatan *Content-Based Filtering* berbasis representasi teks menggunakan TF-IDF dan *cosine similarity* untuk mengukur kesesuaian secara langsung antara CV dan deskripsi pekerjaan. Berbeda dengan penelitian sebelumnya, penelitian ini memanfaatkan data CV dalam format PDF yang diproses melalui tahapan *preprocessing* NLP secara komprehensif. Selain itu, penelitian ini juga mengintegrasikan model ke dalam aplikasi berbasis *Streamlit*, sehingga memungkinkan pengguna memperoleh rekomendasi pekerjaan secara interaktif dan aplikatif.

#### B. Sistem Rekomendasi

Sistem rekomendasi (*Recommender Systems*) telah berkembang sebagai solusi penting untuk mengatasi masalah kelebihan informasi (*information overload*) yang dihadapi pengguna di lingkungan digital. Sistem ini dapat didefinisikan sebagai kelas perangkat lunak yang bertujuan untuk menyaring dan menyajikan item-item (seperti produk, film, atau dalam konteks ini, lowongan pekerjaan) yang paling relevan dan menarik bagi pengguna.

Menurut penelitian [14], terdapat beberapa pendekatan utama yang digunakan untuk menghasilkan prediksi dan rekomendasi. Pendekatan-pendekatan ini secara luas diklasifikasikan ke dalam tiga kategori utama: *Content-Based Filtering* (CBF), *Collaborative Filtering* (CF), dan *Hybrid Approach*.

Penelitian ini berfokus pada *Content-Based Filtering* (CBF). Pendekatan CBF bekerja dengan cara menganalisis fitur atau “konten” dari item itu sendiri. Sistem ini merekomendasikan item-item baru yang memiliki karakteristik serupa dengan item yang telah disukai pengguna di masa lalu, atau yang cocok dengan profil preferensi pengguna yang telah dibangun. Dalam konteks pencarian lowongan kerja, metode CBF diterapkan dengan menganalisis konten dari *Curriculum Vitae* (CV) kandidat dan mencocokkannya dengan konten dari deskripsi lowongan pekerjaan.

### C. *LinkedIn*

LinkedIn adalah *platform* profesional terbesar di dunia yang menyediakan profil pengguna sebagai proxy digital untuk CV, dan postingan lowongan pekerjaan dengan deskripsi tugas dan kualifikasi yang dibutuhkan [15]. Beberapa studi memanfaatkan data LinkedIn untuk membangun model pencocokan pekerjaan. Penelitian [16] menggunakan data publik LinkedIn dan menerapkan *latent semantic analysis* (LSA) untuk merekomendasikan lowongan pekerjaan berdasarkan keterampilan pengguna. Data LinkedIn dianggap representatif karena menyertakan informasi pendidikan, pengalaman kerja, dan keterampilan.

### D. *Curriculum Vitae* (CV)

*Curriculum Vitae* (CV) merupakan dokumen utama yang menilai kualifikasi, pengalaman, dan keahlian pelamar. Dalam sistem rekomendasi pekerjaan, CV berfungsi sebagai input utama untuk analisis konten dalam CBF [17]. Representasi vektor CV berbasis TF-IDF efektif menyoroti kompetensi unik pelamar dan relevan terhadap deskripsi pekerjaan [18].

### E. *Natural Language Processing* (NLP)

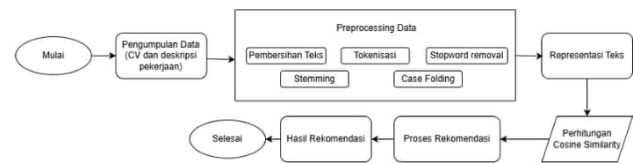
*Natural Language Processing* (NLP) adalah sebuah cabang ilmu komputer yang berkaitan dengan pemrosesan bahasa alami manusia oleh mesin atau komputer [19]. NLP digunakan untuk memahami dan memproses bahasa manusia secara otomatis. Dalam sistem rekomendasi pekerjaan, NLP digunakan untuk preprocessing CV dan deskripsi lowongan, termasuk ekstraksi teks, *case folding*, *stopword removal*, dan *stemming*. Teknik NLP memastikan setiap kata kunci (*keyword*) penting dalam CV atau deskripsi pekerjaan diidentifikasi dan diproses menjadi fitur numerik untuk algoritma CBF [15].

### F. *Content-Based Filtering* (CBF)

*Content-Based Filtering* (CBF) yang diterapkan dalam penelitian ini adalah *unsupervised learning*. *Unsupervised Learning* adalah pendekatan dimana model dilatih menggunakan data yang tidak berlabel (tidak ada variabel target). Tujuan utamanya adalah menemukan struktur tersembunyi, pola, atau kesamaan dalam data [20]. CBF bekerja dengan mengukur kemiripan antara konten item dan profil pengguna. Dalam sistem rekomendasi pekerjaan, konten tersebut adalah teks CV dan deskripsi pekerjaan. Representasi

dilakukan dengan TF-IDF yang kemudian diubah menjadi vektor, dan kesamaan dihitung menggunakan cosine similarity. Dalam konteks *cold start*, model berbasis konten sangat sesuai karena tidak memerlukan riwayat interaksi pengguna [21]. Perhitungan *cosine similarity* menghasilkan skor antara 0 hingga 1 [22].

## II. METODOLOGI PENELITIAN



Gbr 1. Flowchart Penelitian

Penelitian ini menerapkan metodologi *Content-Based Filtering* (CBF) untuk membangun sistem rekomendasi pekerjaan yang sistematis. Pendekatan ini berfokus pada analisis konten tekstual untuk mencocokkan *Curriculum Vitae* (CV) kandidat dengan deskripsi lowongan pekerjaan dari *LinkedIn*.

Alur penelitian, yang ditunjukkan pada Gambar 1, dirancang untuk membangun sistem yang relevan melalui beberapa tahapan inti: pengumpulan data, pra-pemrosesan data, representasi fitur (TF-IDF), dan perhitungan relevansi (*Cosine Similarity*) untuk menghasilkan rekomendasi pekerjaan.

### A. *Pengumpulan Data*

Tahap penelitian ini diawali dengan pengumpulan data yang dilakukan secara manual dari platform jejaring profesional LinkedIn. Pengumpulan data berfokus untuk memperoleh dua set data utama yang tersedia secara publik, yaitu dataset *Curriculum Vitae* (CV) kandidat dan dataset deskripsi lowongan pekerjaan. Seluruh data yang dikumpulkan ini digunakan secara utuh baik sebagai basis data sistem maupun sebagai objek dalam skenario pengujian sistem.

Dataset pertama terdiri dari 106 dokumen *Curriculum Vitae* (CV) dalam format PDF yang diperoleh dari berbagai profil pelamar kerja. Seluruh 106 CV tersebut digunakan sebagai skenario pengujian sistem, di mana setiap CV diproses secara independen untuk menghasilkan rekomendasi pekerjaan berdasarkan tingkat kemiripan terhadap 902 deskripsi lowongan pekerjaan.

Karakteristik CV yang digunakan cukup beragam dan merepresentasikan berbagai profil pencari kerja. CV tersebut berasal dari empat bidang utama yang sesuai dengan kategori lowongan pekerjaan dalam penelitian, yaitu IT dan Data, Keuangan dan Akuntansi, Teknik dan Produksi, serta Bidang Lainnya. Selain memiliki variasi latar belakang profesi, CV yang digunakan juga mencakup perbedaan tingkat pendidikan, pengalaman kerja, keterampilan teknis maupun nonteknis, sertifikasi profesional, serta pengalaman proyek yang berbeda-beda. Keragaman karakteristik tersebut bertujuan untuk memberikan gambaran yang lebih representatif terhadap

kondisi nyata pencari kerja dan memungkinkan evaluasi sistem rekomendasi pada berbagai latar belakang kandidat.

Penelitian ini tidak menerapkan pembagian data *training* dan *testing* sebagaimana pada pendekatan pembelajaran terawasi (*supervised learning*), karena sistem rekomendasi yang dibangun menggunakan metode *Content-Based Filtering* berbasis TF-IDF dan *Cosine Similarity* yang bekerja melalui perhitungan tingkat kemiripan antar dokumen tanpa proses pelatihan model maupun penggunaan data berlabel. Oleh karena itu, seluruh data lowongan pekerjaan digunakan sebagai basis pencarian rekomendasi, sedangkan 106 dokumen CV digunakan sebagai skenario pengujian sistem.

Untuk mempermudah proses ekstraksi informasi dan pencocokan, dokumen CV yang semula berbentuk teks tidak terstruktur direpresentasikan ke dalam lima dimensi utama, yaitu: (1) Profile Summary yang berisi ringkasan profil dan kompetensi pelamar; (2) Skills yang memuat keterampilan teknis maupun nonteknis; (3) Experience yang mencakup pengalaman kerja, proyek, dan tanggung jawab pekerjaan; (4) Education yang berisi riwayat pendidikan formal; serta (5) Certification yang memuat sertifikasi atau pelatihan yang pernah diikuti.

Dataset kedua adalah deskripsi lowongan pekerjaan yang dikumpulkan secara manual dari platform LinkedIn. Dataset ini terdiri dari 902 data lowongan pekerjaan yang dikelompokkan ke dalam empat bidang utama, yaitu:

1. IT dan Data
2. Keuangan dan Akuntansi
3. Teknik dan Produksi
4. Bidang Lainnya

Sebagaimana ditunjukkan pada Tabel II, setiap entri lowongan pekerjaan memuat informasi posisi, perusahaan, lokasi, serta uraian pekerjaan yang menjadi dasar analisis kesesuaian dengan *Curriculum Vitae* (CV) kandidat. Untuk mendukung proses pencocokan, setiap data lowongan direpresentasikan ke dalam empat dimensi utama, yaitu: (1) *Category* yang menunjukkan kategori atau bidang pekerjaan; (2) *Position* yang berisi nama jabatan atau posisi yang ditawarkan; (3) *Required Skills* yang memuat keterampilan yang dipersyaratkan oleh perusahaan; dan (4) *Job Description* yang berisi uraian tugas, tanggung jawab, serta kualifikasi yang dibutuhkan. Struktur tabel ini memperlihatkan keragaman bidang pekerjaan yang digunakan sebagai acuan dalam proses rekomendasi.

Tabel II. DATASET DESKRIPSI ATAU LOWONGAN PEKERJAAN (SUMBER: LINKEDIN)

No.	Position	Company	..	Job description	Job category
1.	Engineering electrical staff	Thamrin Group	..	Membuka lowongan Engineering Electrical Staff ... pengalaman 1-2 tahun.	Teknik dan Produksi

2.	Creative content design junior	Seven Clean Seans	..	Membuka posisi Creative Content & Design Junior di Denpasar. ... Design Junior Application.	Bidang Lainnya
...	...	...	..	...	...
901	Finance accounting staff	PT. Salam Pacific Indonesia Lines	..	Membuka posisi Finance & Accounting Staff di Bali. Membutuhkan ... memimpin tim.	Akuntansi dan Keuangan
902	Backend developer	SoftwareS eni	..	Membutuhkan kandidat dengan pengalaman kuat di backend development, menguasai PHP ... kemampuan tambahan dalam Python.	IT dan Data

Dalam tahap pengujian sistem, seluruh 106 dokumen CV dan 902 data lowongan pekerjaan dilibatkan secara penuh dalam proses perhitungan kemiripan. Setiap CV diproses secara independen untuk menghasilkan rekomendasi pekerjaan berdasarkan nilai *similarity* tertinggi terhadap seluruh lowongan yang tersedia.

Sebagai representasi karakteristik data pengujian, empat contoh CV yang digunakan dalam penelitian ini ditampilkan pada Gbr. 2 hingga Gbr. 5, yaitu "CV\_satu.pdf" yang merepresentasikan bidang IT dan Data, "CV\_dua.pdf" yang merepresentasikan bidang Keuangan dan Akuntansi, "CV\_tiga.pdf" yang merepresentasikan bidang Teknik dan Produksi, serta "CV\_empat.pdf" yang merepresentasikan Bidang Lainnya. Pemilihan empat CV tersebut bertujuan untuk merepresentasikan masing-masing kategori bidang pekerjaan yang digunakan dalam penelitian, sedangkan proses pengujian tetap dilakukan terhadap seluruh 106 CV yang tersedia.

Sistem kemudian melakukan ekstraksi teks menggunakan pustaka PyMuPDF (*fitz*) untuk memperoleh data dalam bentuk teks mentah (*raw text*). Selanjutnya, teks hasil ekstraksi dari 106 CV diproses menggunakan metode TF-IDF dan *Cosine Similarity* untuk menghitung tingkat kemiripan terhadap 902 data lowongan pekerjaan. Hasil perhitungan tersebut digunakan untuk menghasilkan rekomendasi pekerjaan berdasarkan nilai *similarity* tertinggi.



Gbr. 2 CV\_satu.pdf (Bidang IT dan Data)



Gbr. 5 CV\_empat.pdf (Bidang Lainnya)



Gbr. 3 CV\_dua.pdf (Bidang Keuangan dan Akuntansi)



Gbr. 4 CV\_tiga.pdf (Bidang Teknik dan Produksi)

B. Preprocessing Data

Setelah data terkumpul, tahap krusial berikutnya adalah preprocessing data untuk mengubah teks mentah (CV dan lowongan pekerjaan) menjadi format yang bersih dan terstruktur. Proses ini penting untuk mengurangi noise dan menyiapkan data untuk analisis. Berdasarkan alur penelitian (Gambar 1), tahapan ini terdiri dari lima proses inti:

1. **Pembersihan Teks (Text Cleaning):** Menghapus semua karakter yang tidak relevan, seperti tanda baca, angka, dan simbol khusus.
2. **Case Folding:** Menyeragamkan seluruh teks menjadi format huruf kecil (*lowercase*) untuk memastikan konsistensi (seperti: "Admin" dan "admin" dianggap sama).
3. **Tokenisasi (Tokenization):** Memecah kalimat menjadi unit-unit kata (token) individual.
4. **Stopword Removal:** Menyaring dan membuang kata-kata umum yang tidak memiliki makna signifikan dalam pencocokan (seperti "yang", "di", "dan").
5. **Stemming:** Mengubah setiap kata ke bentuk dasarnya (seperti: "mempelajari" atau "dipelajari" menjadi "ajar") agar dapat dianggap sebagai fitur yang sama.

Data bersih yang dihasilkan dari lima langkah ini kemudian siap diproses pada tahap Representasi Teks menggunakan TF-IDF.

C. Representasi Teks

Setelah data bersih dari tahap *preprocessing*, langkah selanjutnya adalah Representasi Teks. Tahap ini bertujuan mengubah data teks (dari CV dan lowongan pekerjaan) menjadi format vektor numerik agar dapat dihitung kemiripannya.

Pada penelitian ini, metode yang digunakan adalah *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF adalah metode statistik yang menghitung bobot sebuah kata berdasarkan dua faktor: frekuensi kemunculannya dalam satu dokumen (*Term Frequency - TF*) dan kelangkaannya di seluruh korpus dokumen (*Inverse Document Frequency - IDF*).

Metode ini menghasilkan representasi vektor yang efektif dalam menangkap kata-kata kunci penting (seperti keterampilan teknis atau jabatan) dari setiap dokumen. Vektor TF-IDF inilah yang kemudian digunakan dalam tahap perhitungan kesamaan [23].

$$W_{(t,d)} = tf_{(t,d)} * idf_t \quad (1)$$

Keterangan:

$tf_{(t,d)}$  : jumlah  $t$  dalam  $d$  / jumlah kata pada  $d$   
 $idf_t$  : jumlah dokumen yang membuat  $term$

#### D. Perhitungan Cosine Similarity

Tahap perhitungan *cosine similarity* adalah inti dari proses pencocokan *Content-Based Filtering*. Setelah CV dan lowongan pekerjaan berhasil direpresentasikan sebagai vektor numerik menggunakan TF-IDF, metode ini diterapkan untuk mengukur seberapa “mirip” kedua vektor tersebut.

*Cosine similarity* menghitung kosinus sudut antara vektor CV dan vektor lowongan. Perhitungan ini menghasilkan skor antara 0 (sangat berbeda) hingga 1 (sangat mirip). Nilai yang lebih tinggi menunjukkan tingkat kesesuaian yang lebih besar antara konten (keterampilan, pengalaman) di CV dengan kriteria yang ada di deskripsi lowongan. Skor inilah yang kemudian digunakan untuk mengurutkan dan menghasilkan rekomendasi [24].

$$CosSim(A, B) = \frac{A \cdot B}{\|A\| \cdot \|B\|} \quad (2)$$

Keterangan:

$A$  dan  $B$  : vektor fitur yang dihasilkan dari representasi TF-IDF dari teks, seperti deskripsi pekerjaan dan profil pengguna.

$A \cdot B$  : hasil perkalian *dot product* antara vektor TF-IDF, mengukur kesamaan berdasarkan kata-kata penting yang sama-sama muncul.

$|A|$  dan  $|B|$  : magnitudo (panjang) dari masing-masing vektor TF-IDF, dihitung dari akar kuadrat jumlah kuadrat elemen vektor, digunakan untuk normalisasi.

#### E. Proses Rekomendasi

Tahap Proses Rekomendasi adalah langkah operasional di mana skor *cosine similarity* (yang diperoleh dari tahap sebelumnya) digunakan untuk menghasilkan daftar peringkat. Sistem akan mengurutkan semua lowongan pekerjaan berdasarkan skor kemiripannya terhadap sebuah CV, dari nilai tertinggi ke terendah. Sejumlah lowongan teratas (Top-5) yang memiliki skor tertinggi kemudian dipilih. Lowongan-lowongan inilah yang dianggap paling relevan dan akan disajikan kepada pengguna pada tahap hasil rekomendasi.

#### F. Hasil Rekomendasi

Tahap hasil rekomendasi adalah keluaran akhir dari sistem. Berdasarkan proses pemeringkatan pada tahap sebelumnya, sistem akan menghasilkan daftar lowongan pekerjaan yang dinilai paling relevan untuk CV kandidat.

Skor *cosine similarity* menentukan urutan daftar ini. Skor yang lebih tinggi menandakan tingkat kesesuaian yang lebih kuat antara keterampilan dan pengalaman di CV dengan kriteria lowongan. Hasil ini kemudian disajikan dalam format daftar terurut yang dapat berisi informasi seperti posisi pekerjaan, nama perusahaan, skor kesesuaian, dan tautan lowongan.

### III. HASIL DAN PEMBAHASAN

Sistem menampilkan lima pekerjaan terbaik (Top-5) berdasarkan nilai similarity tertinggi antara dokumen CV dan deskripsi lowongan pekerjaan. Untuk mengevaluasi kemampuan sistem pada berbagai karakteristik kandidat, penelitian ini menggunakan empat skenario pengujian CV yang mewakili bidang IT dan Data, Keuangan dan Akuntansi, Teknik dan Produksi, serta Bidang Lainnya. Setiap CV diproses secara independen terhadap 902 data lowongan pekerjaan sehingga menghasilkan rekomendasi pekerjaan yang sesuai dengan profil masing-masing kandidat. Empat dokumen CV tersebut dipilih sebagai sampel representatif dari total 106 CV yang digunakan dalam penelitian. Pemilihan dilakukan dengan mempertimbangkan keterwakilan masing-masing kategori bidang pekerjaan, yaitu IT dan Data, Keuangan dan Akuntansi, Teknik dan Produksi, serta Bidang Lainnya.

Meskipun penelitian menggunakan 106 dokumen CV sebagai data pengujian, pada bagian hasil hanya ditampilkan empat skenario CV yang dipilih sebagai sampel representatif dari masing-masing kategori bidang pekerjaan. Pemilihan ini dilakukan untuk memudahkan analisis dan pembahasan hasil rekomendasi yang dihasilkan sistem. Seluruh 106 dokumen CV tetap diproses dalam pengujian sistem untuk menghasilkan rekomendasi pekerjaan. Namun, karena keterbatasan ruang pada artikel, hanya empat CV yang ditampilkan sebagai contoh representatif untuk mewakili masing-masing kategori bidang pekerjaan.

Sebelum menampilkan hasil rekomendasi, penelitian ini menggunakan empat dokumen CV sebagai skenario pengujian yang mewakili masing-masing kategori bidang pekerjaan dalam dataset. Pemilihan skenario ini bertujuan untuk mengevaluasi kemampuan sistem dalam menghasilkan rekomendasi yang relevan pada berbagai karakteristik kandidat. Skenario pengujian yang digunakan ditunjukkan pada Tabel III.

Tabel III. SKENARIO PENGUJIAN CURRICULUM VITAE (CV)

Skenario	Nama CV	Bidang
Skenario 1	CV_satu.pdf	IT dan Data
Skenario 2	CV_dua.pdf	Keuangan dan Akuntansi
Skenario 3	CV_tiga.pdf	Teknik dan Produksi
Skenario 4	CV_empat.pdf	Bidang Lainnya

Berdasarkan Tabel III, setiap skenario pengujian merepresentasikan bidang pekerjaan yang berbeda sehingga dapat digunakan untuk mengamati konsistensi sistem dalam memberikan rekomendasi pekerjaan sesuai dengan karakteristik dan kompetensi kandidat. Hasil rekomendasi untuk masing-masing skenario pengujian selanjutnya disajikan pada Tabel IV hingga Tabel VII.

Hasil rekomendasi untuk skenario pertama, yaitu CV\_satu.pdf yang merepresentasikan bidang IT dan Data, ditampilkan pada Tabel IV. Sistem menghasilkan lima rekomendasi pekerjaan dengan nilai similarity tertinggi berdasarkan proses perhitungan menggunakan metode TF-IDF dan Cosine Similarity.

Tabel IV. HASIL REKOMENDASI PEKERJAAN DARI CONTOH "CV\_SATU.PDF (BIDANG IT DAN DATA)"

Position	Company	Location	Skor Similarity
Backend developer	SoftwareSeni	Yogyakarta	0.666
Backend engineer	Bridestory	Tangerang	0.645
Junior software engineer	Synpulse8	Jakarta	0.623
Software quality assurance	PT. Astra Graphia	Jakarta	0.620
System integration sect head	PT. Suryaraya Rubberindo Industries	Bogor	0.614

Berdasarkan Tabel IV, sistem menghasilkan rekomendasi yang sesuai dengan latar belakang kandidat pada bidang IT dan Data. Empat dari lima rekomendasi teratas berada pada kategori pengembangan perangkat lunak, yaitu *Backend Developer*, *Backend Engineer*, *Junior Software Engineer*, dan *Software Quality Assurance*. Nilai similarity tertinggi sebesar 0,666 pada posisi *Backend Developer* menunjukkan bahwa metode TF-IDF dan *Cosine Similarity* mampu mengidentifikasi kesesuaian kompetensi kandidat dengan kebutuhan lowongan pekerjaan secara efektif. Hasil ini menunjukkan bahwa sistem dapat menghasilkan rekomendasi yang relevan berdasarkan informasi yang terdapat pada CV.

Selanjutnya, Tabel V menampilkan hasil rekomendasi untuk CV\_dua.pdf yang merepresentasikan bidang Keuangan dan Akuntansi. Pengujian ini dilakukan untuk mengetahui kemampuan sistem dalam mengidentifikasi kesesuaian kompetensi kandidat dengan kebutuhan lowongan pada bidang keuangan.

Tabel V. HASIL REKOMENDASI PEKERJAAN DARI CONTOH "CV\_DUA.PDF (BIDANG KEUANGAN DAN AKUNTANSI)"

Position	Company	Location	Skor Similarity
Finance accounting staff	PT Salam Pacific Indonesia Lines	Jakarta	0.602

Finance officer scg indonesia	SCG	Jakarta	0.601
Account controller 2 general accounting account payable	PT. Suryaraya Rubberindo Industries	Bogor	0.588
Electrical engineering antenna rf if batam	PASIFIK SATELIT NUSANTARA	Batam	0.572
Product finance analyst accommodation and lifestyle	Traveloka	Jakarta	0.565

Berdasarkan Tabel V, sistem menghasilkan rekomendasi yang didominasi oleh posisi pada bidang keuangan dan akuntansi, seperti *Finance Accounting Staff*, *Finance Officer*, dan *Account Controller*. Hasil ini menunjukkan bahwa kompetensi yang terdapat pada CV\_dua.pdf memiliki tingkat kesesuaian yang baik dengan kebutuhan lowongan pada bidang tersebut. Nilai similarity tertinggi sebesar 0,602 menunjukkan bahwa metode Content-Based Filtering mampu mengidentifikasi dan merekomendasikan pekerjaan yang relevan berdasarkan informasi yang terdapat pada CV kandidat.

Hasil rekomendasi untuk CV\_tiga.pdf yang mewakili bidang Teknik dan Produksi ditunjukkan pada Tabel VI. Skenario ini digunakan untuk menguji kemampuan sistem dalam mencocokkan keterampilan teknis dan pengalaman kerja kandidat dengan kebutuhan industri teknik dan produksi.

Tabel VI. HASIL REKOMENDASI PEKERJAAN DARI CONTOH "CV\_TIGA.PDF (BIDANG TEKNIK DAN PRODUKSI)"

Position	Company	Location	Skor Similarity
L2 engineer data center facility specialist	bitcorp. (PT. Bringin Inti Teknologi)	Jakarta	0.821
Teknisi mesin produksi jakarta	PT Penerbit Erlangga	Jakarta	0.685
Leader production	PT Mandis Plastic Technology	Tangerang	0.593
Marketing agent officer motorku	WOM Finance	Jakarta	0.583
Leader production	PT Mandis Plastic Technology	Tangerang	0.557

Berdasarkan Tabel VI, sistem memberikan rekomendasi yang didominasi oleh posisi pada bidang teknik dan produksi. Nilai similarity tertinggi sebesar 0,821 menunjukkan tingkat kesesuaian yang tinggi antara kompetensi yang dimiliki kandidat dengan kebutuhan lowongan pekerjaan. Hasil ini mengindikasikan bahwa metode TF-IDF dan *Cosine Similarity* mampu mengidentifikasi keterampilan teknis yang relevan pada CV dan deskripsi pekerjaan.

Tabel VII menyajikan hasil rekomendasi untuk CV\_empat.pdf yang merepresentasikan Bidang Lainnya. Skenario ini digunakan untuk mengevaluasi performa sistem pada kandidat yang berasal dari bidang non-teknis dan non-keuangan.

Tabel VII. HASIL REKOMENDASI PEKERJAAN DARI CONTOH "CV EMPAT.PDF (BIDANG LAINNYA)"

Position	Company	Location	Skor Similarity
HRGA SUPERVISOR	Agrosari Farm	Kupang	0.789
Human Resources Supervisor	SinarAlam Group	Banjarmasin	0.698
Head of Human Capital	Confidential Jobs	Tangerang	0.618
Kepala Personalia (HRD)	PT. Tri Sinar Purnama	Semarang	0.601
L2 Engginer Data Center Facility Specialist	bitcorp. (PT. Bringin Inti Teknologi)	Bali	0.562

Berdasarkan Tabel VII, sistem menghasilkan rekomendasi yang sesuai dengan karakteristik CV pada kategori Bidang Lainnya, yang didominasi oleh posisi *Human Resource* dan *Human Capital*. Mayoritas rekomendasi memiliki nilai similarity di atas 0,6 sehingga menunjukkan tingkat relevansi yang baik antara profil kandidat dan kebutuhan perusahaan.

Untuk memberikan gambaran ringkas mengenai hasil pengujian pada seluruh skenario CV yang ditampilkan, ringkasan kategori rekomendasi dominan dan nilai similarity tertinggi disajikan pada Tabel VIII.

Tabel VIII. RINGKASAN HASIL PENGUJIAN SKENARIO CURRICULUM VITAE (CV)

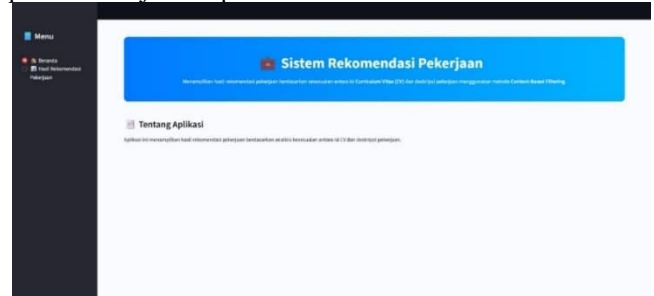
Nama CV	Bidang CV	Rekomendasi Dominan	Similarity Tertinggi
CV_satu	IT dan Data	Software Engineer	0.666
CV_dua	Kuangan dan Akuntansi	Finance & Accounting	0.602
CV_tiga	Teknik dan Produksi	Teknik & Produksi	0.821
CV_empat	Bidang Lainnya	Human Resource	0.789

Berdasarkan Tabel VIII, terlihat bahwa setiap CV menghasilkan pola rekomendasi yang berbeda sesuai dengan bidang yang direpresentasikan. CV bidang IT dan Data memperoleh rekomendasi yang didominasi oleh posisi pengembangan perangkat lunak, sedangkan CV bidang Keuangan dan Akuntansi memperoleh rekomendasi pada posisi *finance* dan *accounting*. Pada CV bidang Teknik dan Produksi, nilai similarity tertinggi mencapai 0,821, yang merupakan nilai tertinggi dibandingkan skenario lainnya. Sementara itu, CV bidang lainnya memperoleh rekomendasi yang didominasi oleh posisi *Human Resource* dan *Human Capital*. Hasil ini menunjukkan bahwa sistem mampu membedakan karakteristik

kandidat dan menghasilkan rekomendasi yang sesuai dengan konten CV yang diberikan.

#### A. Tampilan Antarmuka Sistem

Sistem diimplementasikan menggunakan Streamlit dengan antarmuka sederhana dan interaktif. Pengguna dapat mengunggah CV dalam format PDF, yang kemudian diproses otomatis untuk menampilkan daftar rekomendasi pekerjaan beserta tautan asli menuju lowongan di LinkedIn. Tampilan aplikasi ditunjukkan pada Gbr. 6.



Gbr 6. Tampilan Antarmuka Streamlit Sistem Rekomendasi Pekerjaan

Tampilan ini dirancang agar profesional, mudah digunakan, dan responsif bagi pengguna dengan latar belakang pendidikan yang beragam.

#### B. Evaluasi Sistem Rekomendasi

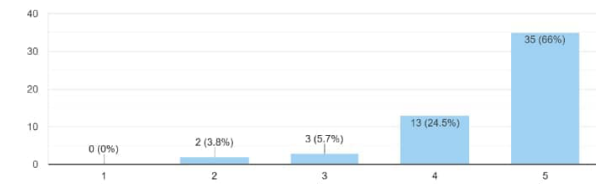
Evaluasi sistem ini menggunakan survei yang ada pada Tabel IV. mendapatkan 53 responden dilakukan untuk menilai relevansi hasil rekomendasi dan kepuasan pengguna terhadap sistem rekomendasi pekerjaan berbasis *Content-Based Filtering* (CBF). Responden berasal dari berbagai latar belakang, seperti pelajar, mahasiswa, *fresh graduate*, dan karyawan. Hasil evaluasi menunjukkan bahwa sistem mampu memberikan rekomendasi yang relevan dengan profil dan pengalaman pengguna.

Tabel IX. RINGKASAN HASIL EVALUASI

Aspek	Hasil
Rata-rata Relevansi Rekomendasi	4.2 / 5
Rata-rata Kepuasan Pengguna	4.5 / 5
Respon Positif	90,5% pengguna puas dan sangat puas
Masukan	Sistem mudah digunakan dan hasilnya relevan

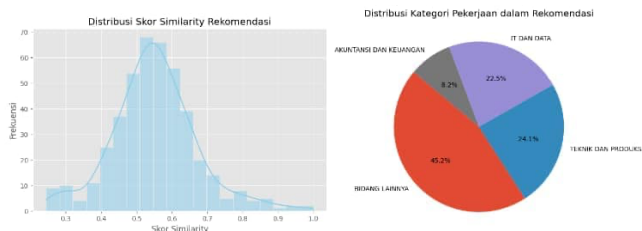
Untuk menilai kepuasan pengguna terhadap sistem rekomendasi yang dibuat, dilakukan survei dengan pertanyaan dan hasilnya ditampilkan pada Gbr 7. berikut.

Secara keseluruhan, seberapa puasakah Anda dengan kualitas rekomendasi yang diberikan oleh sistem?  
53 responses



Gbr 7. Evaluasi Hasil Kualitas Sistem Rekomendasi (Kuisisioner)

Dari data terlihat mayoritas responden memberikan penilaian positif, dengan 90,5% merasa puas hingga sangat puas. Rata-rata skor relevansi 4,2/5 memperkuat temuan bahwa sistem mampu menyajikan rekomendasi sesuai kebutuhan dan isi CV. Selain relevansi, kemudahan penggunaan menjadi nilai tambah karena pengguna cukup mengunggah CV dalam format PDF. Meski demikian, beberapa masukan menyarankan agar tampilan hasil analisis dan rekomendasi dibuat lebih jelas dan terstruktur. Untuk memperkuat temuan ini, penelitian juga menambahkan visualisasi hasil rekomendasi pada Gbr 8. berikut.



Gbr 8. Histogram distribusi skor *similarity* dan *pie chart* kategori pekerjaan

Gbr 8. menampilkan *histogram* distribusi skor *similarity* dan *pie chart* kategori pekerjaan yang memperlihatkan bahwa sebagian besar rekomendasi memiliki skor di atas 0.6, menegaskan relevansi tinggi antara CV dan deskripsi pekerjaan serta menunjukkan yang mengilustrasikan keragaman bidang rekomendasi meliputi IT & Data, Keuangan & Akuntansi, Teknik & Produksi, serta Bidang Lainnya. Kedua visualisasi ini menegaskan bahwa sistem tidak hanya relevan, tetapi juga mampu memberikan cakupan rekomendasi yang luas dan beragam.

Secara keseluruhan, sistem berbasis *Content-Based Filtering* ini menunjukkan efektivitas tinggi dalam mendukung proses pencarian kerja secara praktis dan tepat sasaran. CBF dipilih karena hanya memerlukan konten teks dari CV dan deskripsi lowongan, sehingga tetap dapat bekerja meski tanpa riwayat interaksi pengguna. Pendekatan ini sangat sesuai untuk konteks penelitian berbasis CV PDF.

Sebaliknya, *Collaborative Filtering* (CF) membutuhkan data interaksi pengguna–lowongan dalam jumlah besar sehingga tidak cocok untuk kondisi tanpa histori perilaku. CF juga lemah pada pengguna atau lowongan baru, meskipun berpotensi memberi rekomendasi yang lebih personal bila datanya mencukupi. Untuk penelitian lanjutan, pengumpulan log klik atau lamaran dapat memungkinkan evaluasi empiris antara

CBF dan CF, termasuk eksplorasi pendekatan *hybrid*. Dengan demikian, sistem dapat dikembangkan menjadi lebih adaptif dan akurat di masa depan.

#### IV. KESIMPULAN

Penelitian ini berhasil mengimplementasikan sistem rekomendasi pekerjaan berbasis *Content-Based Filtering* yang mampu mengolah data CV dalam format PDF dan mencocokkannya dengan deskripsi lowongan pekerjaan secara efektif. Penggunaan metode TF-IDF dan cosine similarity terbukti mampu menghasilkan rekomendasi dengan tingkat relevansi yang baik, ditunjukkan oleh mayoritas skor di atas 0,6. Selain itu, hasil evaluasi pengguna menunjukkan tingkat kepuasan yang tinggi, sehingga sistem dinilai mampu membantu proses pencarian kerja secara lebih efisien dan tepat sasaran. Keunggulan sistem terletak pada kemampuannya memanfaatkan data teks tidak terstruktur serta kemudahan penggunaan melalui antarmuka interaktif. Namun, penelitian ini masih memiliki keterbatasan pada jumlah dataset dan belum mengintegrasikan metode berbasis pembelajaran lanjutan. Oleh karena itu, pengembangan selanjutnya disarankan untuk memperluas dataset, menerapkan pendekatan hybrid filtering, serta mengadopsi teknik NLP yang lebih canggih seperti *word embeddings* atau *Sentence-BERT* guna meningkatkan akurasi dan personalisasi sistem.

#### UCAPAN TERIMA KASIH

Penulis menyampaikan rasa terima kasih kepada Universitas Pembangunan Nasional “Veteran” Jawa Timur, khususnya Program Studi Informatika, atas dukungan fasilitas serta bimbingan yang telah diberikan selama proses penyusunan penelitian ini. Ucapan terima kasih juga ditujukan kepada dosen pembimbing atas arahan dan masukan yang sangat berarti sepanjang pelaksanaan penelitian, serta kepada rekan-rekan yang turut berkontribusi dalam proses pengujian dan validasi model.

Selain itu, penulis juga menyampaikan apresiasi kepada panitia Seminar Nasional Informatika Bela Negara (SANTIKA) 2026 atas kesempatan yang diberikan untuk mempublikasikan hasil penelitian ini.

#### REFERENSI

- [1] LinkedIn, “LinkedIn Work Change Report 2025,” LinkedIn News, 2025. [Online]. Available: <https://news.linkedin.com/2025/work-change-report>
- [2] M. Bone, E. Ehlinger, and F. Stephany, “Skills or Degree? The Rise of Skill-Based Hiring for AI and Green Jobs,” arXiv preprint arXiv:2312.11942, Dec. 2023. [Online]. Available: <https://arxiv.org/abs/2312.11942>
- [3] A. Bauer, “Accounting for Qualification in Mismatch Unemployment,” *Journal for Labour Market Research*, vol. 58, no. 1, Art. no. 27, 2024, doi: 10.1186/s12651-024-00386-7.
- [4] L. A. Flórez and L. Gómez, “The Impact of Skill Mismatch on Unemployment, Informality, and Labour Turnover,” *The Economic and Labour Relations Review*, vol. 35, no. 4, pp. 1–20, 2024, doi: 10.1017/ELR.2024.48.

- [5] G. Montt, "Field-of-Study Mismatch and Overqualification: Labour Market Correlates and Their Wage Penalty," *IZA Journal of Labor Economics*, vol. 6, no. 1, Art. no. 2, 2017, doi: 10.1186/s40172-016-0052-x.
- [6] M. Beręsewicz, H. Cherniaev, A. Mantaj, and R. Pater, "Text Analysis of Job Offers for Mismatch of Educational Characteristics to Labour Market Demands," *Journal of Big Data*, vol. 58, pp. 1799–1825, 2024, doi: 10.1007/s11135-023-01707-7.
- [7] D. Ç. Ertuğrul and S. Bitirim, "Job Recommender Systems: A Systematic Literature Review, Applications, Open Issues, and Challenges," *Journal of Big Data*, vol. 12, Art. no. 140, 2025, doi: 10.1186/s40537-025-01173-y.
- [8] J. F. Levid, D. Wijaya, H. Irsyad, dan A. Rahman, "Penerapan Smart, EDAS, dan Cosine Similarity dalam rekomendasi lowongan pekerjaan di era digital," *Buletin Ilmiah Informatika Teknologi*, vol. 3, no. 3, hlm. 85–92, 2025, doi: 10.58369/biit.v2i3.128.
- [9] P. Singla dan V. Verma, "An Intelligent Job Recommendation System based on Semantic Embeddings and Machine Learning," *Journal of Information Systems Engineering & Management*, vol. 10, no. 5s, pp. 520–542, Jan. 2025, doi: 10.52783/jisem.v10i5s.681.
- [10] N. Reimers and I. Gurevych, "Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online, Nov. 2020, pp. 4512–4525. doi: 10.18653/v1/2020.emnlp-main.365.
- [11] A. Ro'uf, Y. M. Pranoto, and E. Setyati, "Sistem Rekomendasi Pekerjaan Menggunakan Content Based Similarity," *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 12, no. 2, pp. 618–626, Aug. 2023. DOI: <https://doi.org/10.33480/inti.v20i1.6786>
- [12] A. Fitria, S. Zaman, dan M. A. Yaqin, "Sistem Rekomendasi Lowongan Pekerjaan Menggunakan Content-based filtering," *Jurnal Edukasi dan Penelitian Informatika*, vol. 10, no. 3, Desember 2024.
- [13] C. Koloman, R. Maulana, R. D. Z. Putri, dan W. A. Harahap, "Sistem rekomendasi pekerjaan di bidang IT menggunakan algoritma content-based filtering," *Journal of Creative Student Research*, vol. 1, no. 6, hlm. 78–88, Des. 2023, doi: 10.55606/jcsrpolitama.v1i6.2992.
- [14] F. Ricci, L. Rokach, dan B. Shapira, "Introduction to recommender systems handbook," dalam *Recommender Systems Handbook*, Boston, MA: Springer, hlm. 1–35, Okt. 2010, doi: 10.1007/978-0-387-85820-3\_1.
- [15] N. Kumar, M. Gupta, D. Sharma, dan I. Ofori, "Technical job recommendation system using APIs and web crawling," *Computational Intelligence and Neuroscience*, vol. 2022, hlm. 1–11, 2022, doi: 10.1155/2022/7797548.
- [16] Domeniconi, et al., "Job Recommendation From Semantic Similarity of LinkedIn Users' Skills," in *Proc. 5th Int. Conf. on Pattern Recognition Applications and Methods (ICPRAM)*, 2016, pp. 270–277. doi: 10.5220/0005702302700277.
- [17] A. R. Hawari, Munawir, Zulfan, D. Satria, dan Susmanto, "Job recommendation system using the content-based filtering method," *Proceedings of International Conference on Multidisciplinary Research*, vol. 5, no. 2, hlm. 138–140, Des. 2022, doi: 10.32672/pic-mr.v5i2.5419.
- [18] R. Huang, "Improved content recommendation algorithm integrating semantic information," *Journal of Big Data*, vol. 10, Art. no. 84, 2023, doi: 10.1186/s40537-023-00776-7.
- [19] A. Muhaimin, I. Taufik, and D. Daniswara, "Pendeteksian Spam pada E-mail menggunakan Pendekatan Natural Language Processing," *PROSIDING SEMINAR NASIONAL SAINS DATA*, vol. 3, no. 1, pp. 116–121, Nov. 2023.
- [20] D. A. Prasetya, A. Choiruddin, S. Rakshit, A. Puspita Sari, S. S. M. Wara, A. R. Pratama, dan M. Nasrudin, *Machine Learning Praktis dengan Python: Dari Data Mining hingga Neural Network*, Gresik, Jawa Timur: Thalibul Ilmi Publishing and Education, 2025, ISBN: 978-634-7158-90-1.
- [21] A. Singh, V. Pudi, and R. Singh, "Content-based models for cold-start recommendation," *PeerJ Computer Science*, vol. 7, no. e497, 2021. doi: 10.7717/peerj-cs.497.
- [22] T. V. Yadalam, V. M. Gowda, V. S. Kumar, D. Girish, and N. M., "Career Recommendation Systems using Content-based Filtering," in *Proc. 2020 5th Int. Conf. on Communication and Electronics Systems (ICCES)*, Jun. 10–12, 2020, pp. 270–277. doi: 10.1109/ICCES48766.2020.9137992.
- [23] V. M. Hersianty, E. L. Amalia, D. Puspitasari, dan D. W. Wibowo, "Penerapan algoritma TF-IDF dan cosine similarity dalam sistem rekomendasi lowongan pekerjaan," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 1, hlm. 1619–1625, Feb. 2025, doi: 10.36040/jati.v9i1.12406.
- [24] P. Singla dan V. Verma, "An Intelligent Job Recommendation System based on Semantic Embeddings and Machine Learning," *Journal of Information Systems Engineering & Management*, vol. 10, no. 5s, pp. 520–542, Jan. 2025, doi: 10.52783/jisem.v10i5s.681.