

Penerapan Metode *Content-Based Filtering* Dalam Memberikan Rekomendasi Lowongan Kerja Berdasarkan Kualifikasi

Asmara Nindyah Rahati^{1*}, Afifah Mutmainah², Bagas Satrio Ardianto³, Fawwaz Gibran Sagita Al Firdauss⁴, Miftakhu Subkhan⁵

¹Teknik Informatika/Fakultas Ilmu Komputer

Afiliasi Universitas Duta Bangsa Surakarta

^{1*}230103011@mhs.udb.ac.id (penulis korespondensi)

²Teknik Informatika /Fakultas Ilmu Komputer

Afiliasi Universitas Duta Bangsa Surakarta

²230103003@mhs.udb.ac.id

³Teknik Informatika /Fakultas Ilmu Komputer

Afiliasi Universitas Duta Bangsa Surakarta

³230103013@mhs.udb.ac.id

⁴Teknik Informatika /Fakultas Ilmu Komputer

Afiliasi Universitas Duta Bangsa Surakarta

⁴230103021@mhs.udb.ac.id

⁵Teknik Informatika /Fakultas Ilmu Komputer

Afiliasi Universitas Duta Bangsa Surakarta

⁵230103027@mhs.udb.ac.id

Abstrak— Perkembangan platform pencarian kerja digital seperti Glints menghadirkan tantangan baru berupa kelebihan beban informasi (*information overload*) bagi pencari kerja dalam menyeleksi lowongan yang sesuai dengan kualifikasinya. Penelitian ini menerapkan metode *Content-Based Filtering* berbasis *Vector Space Model* (VSM) dengan representasi vektor biner dan perhitungan *Cosine Similarity* untuk mencocokkan profil pelamar dengan syarat lowongan kerja. Sistem diuji menggunakan 20 sampel data lowongan yang diekstraksi dari platform Glints, dengan lima atribut pembanding utama: kategori pekerjaan, posisi, pendidikan, tipe kerja, dan gaji. Hasil pengujian menunjukkan sistem mampu melakukan perankingan lowongan secara *descending* berdasarkan skor kemiripan, menghasilkan tiga kelompok skor (0,7746; 0,6325; dan 0,4472), dengan lowongan *Frontend Developer Internship* dan *Backend Developer Internship* menempati rekomendasi teratas karena kecocokan pada tiga dari lima atribut. Pendekatan ini terbukti sederhana secara komputasi dan mudah diinterpretasikan, namun memiliki keterbatasan berupa nilai kemiripan yang identik antar-lowongan (*tie score*) dan kegagalan pencocokan posisi akibat logika *exact match* yang kurang fleksibel terhadap variasi penamaan posisi di dunia nyata.

Kata kunci— Content-Based Filtering, Cosine Similarity, Sistem Rekomendasi, Lowongan Kerja

Abstract— The development of digital job-search platforms such as Glints has introduced a new challenge in the form of information overload, making it difficult for job seekers to filter vacancies that match their qualifications. This study applies a Content-Based Filtering method based on the Vector Space Model (VSM), using binary vector representation and Cosine Similarity to match applicant profiles with job requirements. The system was tested on 20 job vacancy samples extracted from the Glints platform, using five main comparison attributes: job category, position, education, work type, and salary. The results show that the system successfully ranks vacancies in descending order based on similarity scores, producing three score clusters (0.7746, 0.6325, and 0.4472), with the Frontend Developer Internship and Backend Developer Internship listings ranked highest due to matches on three of the five attributes. This approach proves computationally simple and easy to interpret, but shows limitations in the form of identical similarity scores between vacancies (*tie scores*) and failed position matching caused by an inflexible exact-match logic against real-world position naming variations.

Keywords— Content-Based Filtering, Cosine Similarity, Recommendation System, Job Vacancy

I. PENDAHULUAN

Perkembangan pesat industri teknologi informasi dan digitalisasi di Indonesia telah membawa perubahan signifikan terhadap ekosistem ketenagakerjaan. Seiring dengan tingginya angka angkatan kerja produktif setiap tahunnya, persaingan dalam mencari pekerjaan menjadi semakin kompetitif. Untuk menjembatani antara perusahaan

dan calon pekerja, berbagai platform pencarian kerja digital, seperti Glints, hadir dan berkembang pesat menjadi pilihan utama masyarakat. Platform-platform ini menyediakan ribuan peluang karir yang dapat diakses dari mana saja. Namun, kemudahan akses ini justru melahirkan tantangan baru di tengah masyarakat pencari kerja di Indonesia.

Walaupun ketersediaan informasi lowongan semakin terbuka lebar, kandidat sering kali dihadapkan pada masalah kelebihan beban informasi (*information overload*). Ratusan hingga ribuan lowongan baru muncul setiap harinya dengan kualifikasi yang sangat bervariasi. Hal ini menyebabkan ketidakefisienan waktu dan tenaga, karena kandidat harus menyeleksi ratusan lowongan secara manual untuk menemukan posisi yang benar-benar sesuai dengan latar belakang pendidikan, keahlian, dan preferensi mereka. Kondisi pencarian manual ini tidak jarang membuat pencari kerja merasa kelelahan, yang sering berujung pada pengiriman lamaran yang tidak terarah (*spamming lamaran*), sehingga menurunkan peluang mereka untuk dipanggil oleh rekruter.

Oleh karena itu, diperlukan sebuah solusi teknologi cerdas yang mampu mengotomatisasi penyaringan informasi agar proses pencarian kerja menjadi lebih efektif. Sistem rekomendasi hadir sebagai solusi tepat guna untuk mengatasi permasalahan tersebut dengan memfilter informasi berdasarkan tingkat relevansi pengguna. Terdapat berbagai metode dalam sistem rekomendasi, salah satunya yang paling sesuai untuk mencocokkan kualifikasi adalah *Content-Based Filtering*. Pendekatan ini terbukti efektif dalam merekomendasikan pekerjaan di bidang IT dengan mencocokkan keterampilan pengguna terhadap deskripsi pekerjaan [1]. Secara umum, penerapan *Content-Based Filtering* dapat dilakukan melalui berbagai pendekatan, antara lain pembobotan berbasis teks menggunakan TF-IDF [2], representasi vektor biner pada atribut kategoris [3], serta pengukuran jarak menggunakan *Euclidean Distance* [4]. Masing-masing pendekatan tersebut telah menunjukkan hasil yang positif dalam memberikan rekomendasi pekerjaan yang relevan.

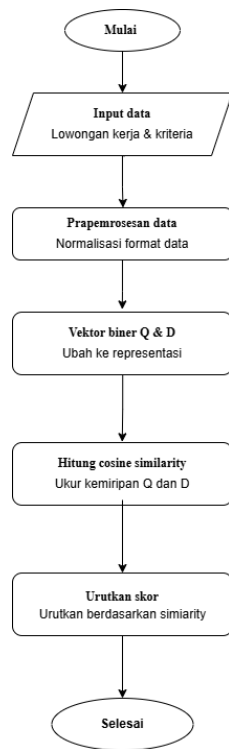
Sejalan dengan penelitian-penelitian tersebut, melalui penelitian ini, diharapkan luaran yang dihasilkan tidak hanya sebatas pembuktian komputasi matematis, melainkan sebuah landasan logika sistem yang dapat diimplementasikan secara nyata pada platform pencarian kerja berskala kecil maupun menengah. Diketahui bahwa pendekatan *Content-Based Filtering* dengan algoritma *MLP Classifier* mampu memberikan rekomendasi pekerjaan dengan akurasi hingga 81% [5]. Hasil dari

pengujian perbandingan ini diharapkan mampu memberikan wawasan (*insight*) yang komprehensif bagi pengembang sistem rekrutmen mengenai pentingnya personalisasi rekomendasi berbasis pembobotan prioritas. Pendekatan serupa yang menggunakan *Cosine Similarity* untuk mencocokkan profil pencari kerja dengan lowongan pekerjaan telah terbukti mencapai akurasi yang baik, yaitu sebesar 80% [2]. Pada akhirnya, penerapan metode ini bertujuan untuk menciptakan pengalaman pencarian kerja yang lebih intuitif, mempercepat proses penyeleksian informasi, dan secara signifikan meningkatkan peluang keberhasilan pelamar dalam mendapatkan pekerjaan yang benar-benar selaras dengan potensi dan ekspektasi mereka.

Penelitian ini mengusulkan penerapan metode *Content-Based Filtering* menggunakan representasi *Vektor Biner* (Binary Vector) dan perhitungan *Cosine Similarity*. Sistem ini dirancang untuk mencocokkan profil pelamar dengan kualifikasi lowongan pekerjaan secara otomatis dan terarah. Untuk menguji efektivitas dan kelayakan metode, digunakan 20 sampel data lowongan kerja aktual yang diekstraksi dari platform Glints. Pengukuran kecocokan kandidat dilakukan dengan memvektorisasi lima atribut utama yang krusial dalam dunia kerja saat ini, yaitu: kategori pekerjaan, posisi yang dilamar, tingkat pendidikan, tipe pekerjaan (*Work From Home/Onsite*), dan penawaran gaji. Pendekatan metrik *Cosine Similarity* ini dipilih karena menawarkan kesederhanaan komputasi melalui pencocokan logika pasti (*exact match*), namun tetap memiliki tingkat akurasi matematis yang tinggi dalam menentukan tingkat relevansi kandidat terhadap syarat operasional lowongan.

II. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan analisis kuantitatif dengan mengimplementasikan metode *Content-Based Filtering* berbasis *Vector Space Model* (VSM).



Gambar 1. Tahapan Penelitian

Model ini memetakan kriteria pelamar dan dokumen lowongan kerja menjadi representasi vektor untuk dihitung tingkat kemiripannya. Tahapan pemodelan sistem terdiri dari langkah-langkah berikut:

1. **Input Data Lowongan Kerja & Kriteria**
Data sekunder yang digunakan merupakan 20 sampel lowongan pekerjaan yang diekstraksi dari platform Glints.[6] Parameter perbandingan difokuskan pada lima atribut utama: Kategori Pekerjaan, Posisi, Pendidikan, Tipe Kerja, dan Gaji.
2. **Prapemrosesan Data (Normalisasi Format Data)**
Tahap prapemrosesan dilakukan untuk menyeragamkan format data. Proses ini meliputi penormalan format seperti mengonversi data string pada penawaran gaji (contoh: "Rp4.000.000") menjadi data numerik murni agar dapat dibandingkan secara matematis pada tahap berikutnya.
3. **Vektor Biner Q & D . (Ubah ke Representasi)**
Sistem menggunakan pendekatan vektor biner untuk mengekstraksi nilai dari atribut lowongan dan kriteria pelamar.[7] Profil

kandidat direpresentasikan sebagai Vektor Kueri Q sedangkan setiap lowongan kerja direpresentasikan sebagai Vektor Dokumen D . Evaluasi fitur menggunakan logika kepastian (*exact match*), di mana setiap dimensi atribut akan bernilai 1 jika kriteria terpenuhi, dan bernilai 0 jika tidak terpenuhi.

4. **Hitung *Cosine Similarity* (Ukur Kemiripan Q dan D)**

Pencocokan antara profil kandidat dengan lowongan kerja dilakukan dengan menghitung jarak sudut antara vektor $||Q||$ dan vektor $||D||$. Metrik ini mengukur kemiripan berdasarkan orientasi vektor, di mana dua vektor yang memiliki arah yang sama ($\cosine = 1$) mengindikasikan dokumen yang sangat mirip, sedangkan vektor yang saling tegak lurus ($\cosine = 0$) mengindikasikan tidak ada kemiripan [8]. Tingkat relevansi (*similarity*) dihitung menggunakan metrik *Cosine Similarity* dengan pemodelan matematis sebagai berikut:

$$\text{Cosine Similarity}(Q, D) = \frac{Q \cdot D}{||Q|| ||D||}$$

Keterangan:

- Q : Representasi vektor biner dari kriteria pelamar.
- D : Representasi vektor biner dari syarat lowongan pekerjaan.
- $Q \cdot D$: Hasil kali titik (dot product) dari kedua vektor.
- $||Q|| ||D||$: Perkalian panjang (magnitudo) dari vektor Q dan D .

5. **Urutkan Skor Berdasarkan *Similarity***

Metrik *Cosine Similarity* menghasilkan nilai pada rentang 0 hingga 1.[9] Nilai yang mendekati 1 mengindikasikan kecocokan yang tinggi antara kualifikasi pelamar dengan syarat lowongan. Sistem kemudian melakukan perangkingan nilai kemiripan secara menurun (*descending*) untuk memberikan daftar rekomendasi pekerjaan yang paling relevan.

III. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil implementasi dan pengujian sistem rekomendasi lowongan kerja menggunakan metode *Content-Based Filtering* yang telah dijelaskan sebelumnya, beserta pembahasan atas temuan-temuan yang diperoleh.

A. Input Data Lowongan Kerja & Kriteria

Pengujian sistem dilakukan terhadap 20 sampel data lowongan kerja yang diekstraksi dari platform Glints, dengan atribut *id_loker*, *posisi*, *kategori*, *tipe_kerja*, *pendidikan*, dan *gaji*. Tabel I menampilkan lima data pertama sebagai contoh.

TABEL I. DATA LOWONGAN KERJA (5 DATA PERTAMA)

id_loker	Posisi	Kategori	Tipe Kerja	Pendidikan	Gaji
1	Graphic Designer	Desain & Kreatif	Freelance	Minimal SMA/SMK	Rp2.000.000
2	KOL Specialist	Pemasaran & Konten	Fulltime	Minimal SMA/SMK	Rp4.300.000
3	Sales & Marketing Staff	Pemasaran & Konten	Fulltime	Minimal Diploma	Rp3.500.000
4	Brand Business Dev. & Growth Strategist	Pemasaran & Konten	Fulltime	Minimal S1	Rp4.000.000
5	Video Editor	Desain & Kreatif	Fulltime	Minimal S1	Rp4.000.000

Sebagai studi kasus, digunakan satu profil pelamar dengan kriteria pencarian sebagaimana ditunjukkan pada Tabel II. Kriteria ini menjadi acuan Vektor Kueri (Q) yang dibandingkan terhadap Vektor Dokumen (D) setiap lowongan.

TABEL II. KRITERIA PELAMAR PADA SKENARIO PENGUJIAN

Atribut	Nilai Kriteria Pelamar
Kategori	Teknologi & Informasi
Posisi	Frontend Developer
Pendidikan	Minimal SMA/SMK
Tipe Kerja	Magang
Gaji Minimal	Rp2.300.000

B. Prapemrosesan Data (Normalisasi Format Data)

Tahap prapemrosesan dilakukan untuk menyeragamkan format data mentah hasil ekstraksi

dari platform Glints. Proses ini berfokus pada pengondisian data penawaran gaji yang awalnya berbentuk teks atau string (seperti contoh pada Tabel I: "Rp4.300.000", "Rp3.500.000") dikonversi menjadi bentuk data numerik murni (4300000, 3500000). Hal ini diperlukan agar sistem dapat melakukan seleksi perbandingan logika nilai matematika terhadap batas gaji minimal yang diinginkan oleh pelamar (Rp2.300.000).

C. Vektor Biner Q & D (Ubah ke Representasi)

Merujuk pada tahapan *binary vectorization*. Setiap dimensi atribut lowongan diberi nilai 1 jika cocok dengan kriteria pelamar, dan bernilai 0 jika tidak cocok.[7] Karena pelamar memasukkan 5 parameter kriteria tetap, maka nilai Vektor Kueri diatur menjadi $Q = [1,1,1,1,1]$, skor similarity yang diperoleh bergantung sepenuhnya pada jumlah dimensi yang bernilai 1 pada vektor D . Tabel III menunjukkan contoh representasi vektor untuk lima data pertama.

TABEL III. CONTOH REPRESENTASI VEKTOR BINER DOKUMEN LOWONGAN

id_loker	Posisi	Vektor D	Keterangan
1	Graphic Designer	[0,0,1,0,0]	Hanya pendidikan cocok
2	KOL Specialist	[0,0,1,0,1]	Pendidikan & gaji cocok
3	Sales & Marketing Staff	[0,0,0,0,1]	Hanya gaji cocok
4	Brand Business Dev. & Growth Strategist	[0,0,0,0,1]	Hanya gaji cocok
5	Video Editor	[0,0,0,0,1]	Hanya gaji cocok

Dari 20 data lowongan yang diuji, sistem menghasilkan tiga kelompok skor kemiripan: 0,7746 (2 lowongan), 0,6325 (7 lowongan), dan 0,4472 (11 lowongan). Lowongan Frontend Developer Internship dan Backend Developer Internship menjadi rekomendasi teratas karena cocok pada tiga dari lima atribut yang dibandingkan, yaitu kategori, pendidikan, dan tipe kerja.

D. Hitung Cosine Similarity (Ukur Kemiripan Q dan D)

Proses pencocokan nilai kemiripan dilakukan dengan menghitung dot product serta perkalian magnitudo antara Vektor Kueri (Q) dan Vektor Dokumen (D) dari 20 data lowongan kerja.[10] Berdasarkan rumus matematis Cosine Similarity,

proses pengujian ini menghasilkan tiga kelompok pengelompokan skor kemiripan yang spesifik pada data lowongan, yaitu nilai skor 0,7746 (terdapat pada 2 lowongan), nilai skor 0,6325 (terdapat pada 7 lowongan), dan nilai skor 0,4472 (terdapat pada 11 lowongan).

E. Urutkan Skor Berdasarkan Similarity

Setelah seluruh nilai skor *Cosine Similarity* didapatkan untuk 20 data lowongan, sistem melakukan pengurutan (*sorting*) secara *descending* sehingga lowongan dengan kecocokan tertinggi ditampilkan sebagai rekomendasi utama. Hasil lengkap pengurutan data untuk menentukan prioritas rekomendasi pekerjaan ditampilkan pada Tabel IV.

TABEL IV. HASIL PERANGKINGAN REKOMENDASI LOWONGAN KERJA (20 DATA)

R an k	I D	Posisi	Kateg ori	Tipe Kerj a	Pendi dikan	Gaji	Cos. Sim.
1	7	Frontend Developer Internship	Teknologi & Informasi	Magang	Min. SMA/SMK	Rp1.000.000	0,7746
1	8	Backend Developer Internship	Teknologi & Informasi	Magang	Min. SMA/SMK	Rp1.000.000	0,7746
3	2	KOL Specialist	Pemasaran & Konten	Fulltime	Min. SMA/SMK	Rp4.300.000	0,6325
3	6	Motion Graphic Designer	Desain & Kreatif	Fulltime	Min. SMA/SMK	Rp5.300.000	0,6325
3	9	Software Developer	Teknologi & Informasi	Fulltime	Min. Diploma	Rp5.000.000	0,6325
3	10	Full Stack Developer - Junior	Teknologi & Informasi	Fulltime	Min. Diploma	Rp4.590.000	0,6325
3	11	AI Developer	Teknologi & Informasi	Kontak	Min. S1	Rp1.200.000	0,6325
3	12	Illustrator	Desain & Kreatif	Fulltime	Min. SMA/SMK	Rp2.500.000	0,6325
3	14	SEO Specialist	Pemasaran & Konten	Fulltime	Min. SMA/SMK	Rp4.000.000	0,6325
10	1	Graphic Designer	Desain & Kreatif	Freelance	Min. SMA/SMK	Rp2.000.000	0,4472
10	3	Sales & Marketing Staff	Pemasaran & Konten	Fulltime	Min. Diploma	Rp3.500.000	0,4472

10	4	Brand Business Dev. & Growth Strategist	Pemasaran & Konten	Fulltime	Min. S1	Rp4.000.000	0,4472
10	5	Video Editor	Desain & Kreatif	Fulltime	Min. S1	Rp4.000.000	0,4472
10	13	Designer Interior	Desain & Kreatif	Freelance	Min. S1	Rp3.000.000	0,4472
10	15	SEO Specialist Intern (Unpaid)	Pemasaran & Konten	Magang	Min. S1	Rp0	0,4472
10	16	Staff Administrator	Administrasi & Operasional	Fulltime	Min. Diploma	Rp4.000.000	0,4472
10	17	Admin Project	Administrasi & Operasional	Fulltime	Min. S1	Rp5.000.000	0,4472
10	18	Manajer Operasional	Administrasi & Operasional	Fulltime	Min. S1	Rp7.000.000	0,4472
10	19	HR Supervisor	Administrasi & Operasional	Fulltime	Min. S1	Rp4.700.000	0,4472
10	20	Staff Accounting	Administrasi & Operasional	Fulltime	Min. S1	Rp4.000.000	0,4472

Berdasarkan hasil pengurutan pada Tabel IV, sistem mengambil dua lowongan dengan posisi *Frontend Developer Internship* dan *Backend Developer Internship* pada peringkat teratas (Rank 1) karena memiliki skor kemiripan tertinggi (0,7746). Hal tersebut disebabkan adanya kecocokan pada tiga dari lima atribut kriteria yang dibandingkan, yaitu kategori, pendidikan, dan tipe kerja.

F. Pembahasan

Interpretasi Hasil Rekomendasi. Sistem berhasil menempatkan dua lowongan magang di bidang Teknologi & Informasi sebagai rekomendasi teratas, sesuai dengan kriteria pelamar yang mencari posisi Frontend Developer di bidang IT dengan tipe kerja Magang. Hal ini menunjukkan pendekatan vektor biner mampu menangkap kecocokan pada atribut kategoris (kategori, pendidikan, tipe kerja) dengan cukup baik.

Analisis Nilai Kemiripan yang Identik (Tie Score). Ditemukan banyak lowongan dengan skor identik (misalnya 7 lowongan bernilai 0,6325 persis sama), karena representasi vektor biner dengan hanya 5 atribut menghasilkan ruang nilai kemiripan yang terbatas secara matematis hanya ada 6 kemungkinan skor berbeda, bergantung jumlah atribut yang cocok. Akibatnya, lowongan yang secara substansi berbeda (misalnya Software Developer dan Illustrator) dapat memperoleh peringkat setara. Ini menjadi keterbatasan utama pendekatan vektor biner dibandingkan pembobotan berbasis teks seperti TF-IDF yang menghasilkan nilai kemiripan lebih bervariasi (kontinu).

Keterbatasan Pencocokan Posisi (Exact Match). Atribut posisi pada seluruh 20 data pengujian selalu bernilai 0 karena logika pencocokan bersifat exact match. Kriteria pelamar mencari “Frontend Developer”, sedangkan data hanya memiliki “Frontend Developer Internship” secara semantik berkaitan erat, namun karena perbandingan dilakukan secara harfiah, atribut posisi tidak pernah berkontribusi pada skor akhir. Hal ini menunjukkan pendekatan exact match kurang fleksibel untuk data lowongan kerja dunia nyata yang memiliki variasi penamaan posisi.

Kelebihan Metode. Pendekatan *Content-Based Filtering* berbasis vektor biner dan *Cosine Similarity* ini memiliki dua kelebihan utama. Pertama, dari segi komputasi, metode ini sederhana dan mudah diinterpretasikan karena setiap skor kemiripan dapat dikaitkan langsung dengan atribut yang cocok maupun tidak. *Cosine Similarity* mengukur sudut antar vektor dalam ruang multidimensi, di mana semakin kecil sudut yang terbentuk menunjukkan tingkat kemiripan yang semakin tinggi, tanpa terpengaruh oleh panjang pendeknya dokumen [8]. Kedua, metode ini tidak memerlukan data interaksi pengguna lain, sehingga mampu mengatasi permasalahan *cold start* kondisi di mana sistem kesulitan memberikan rekomendasi akibat keterbatasan data historis pengguna [3], [11]. Dengan demikian, sistem dapat langsung memberikan rekomendasi hanya berdasarkan satu profil kriteria pelamar, tanpa bergantung pada riwayat interaksi yang seringkali menjadi kelemahan metode *collaborative filtering*.

Saran Perbaikan. Berdasarkan temuan di atas, beberapa perbaikan yang dapat dipertimbangkan untuk pengembangan sistem selanjutnya:

- 1 Mengganti pencocokan posisi dari exact match menjadi pencocokan berbasis kemiripan teks (misalnya TF-IDF atau substring matching).
- 2 Memberikan bobot berbeda pada tiap atribut (weighted cosine similarity), karena tidak semua atribut memiliki tingkat kepentingan yang sama. Pendekatan normalisasi dapat menjadi referensi, di mana bobot kata yang sering muncul di banyak dokumen dikurangi nilainya sehingga kata-kata yang lebih unik dan spesifik memiliki kontribusi lebih besar dalam perhitungan kemiripan [8].
- 3 Menambah jumlah atribut pembanding untuk memperluas ruang nilai kemiripan sehingga mengurangi skor yang identik antar-lowongan.
- 4 Melakukan pengujian akurasi kuantitatif (precision, recall, F1-score) terhadap data acuan (ground truth) untuk mengukur relevansi rekomendasi secara objektif.

IV. KESIMPULAN

Penelitian ini berhasil membuktikan bahwa penerapan metode *Content-Based Filtering* berbasis *Vector Space Model* (VSM) dengan representasi vektor biner dan metrik *Cosine Similarity* dapat digunakan sebagai landasan logika sistem rekomendasi lowongan kerja. Berdasarkan hasil pengujian terhadap 20 sampel data aktual dari platform Glints, sistem mampu mencapai tujuan utama penelitian dengan melakukan penyaringan otomatis dan perangkingan lowongan secara terarah berdasarkan kualifikasi pelamar. Melalui skenario pengujian yang diterapkan, sistem berhasil mengelompokkan skor kemiripan ke dalam tiga variasi nilai spesifik serta menempatkan posisi lowongan magang di bidang teknologi informasi (*Frontend Developer Internship* dan *Backend Developer Internship*) pada prioritas rekomendasi teratas karena kecocokan atribut kategoris yang diuji. Selain memiliki keunggulan berupa kesederhanaan komputasi yang mudah diinterpretasikan, metode ini terbukti efektif dalam mengatasi permasalahan *cold start* karena proses pemfilteran dapat langsung berjalan secara mandiri hanya dengan mengandalkan satu data profil kriteria

pelamar tanpa membutuhkan riwayat interaksi pengguna lain.

Meskipun logika sistem berjalan dengan baik secara matematis, penelitian ini juga mengungkap beberapa keterbatasan utama yang krusial untuk dikembangkan pada model selanjutnya. Pendekatan representasi vektor biner dengan jumlah atribut yang terbatas menyebabkan munculnya fenomena nilai kemiripan yang identik (*tie score*) untuk beberapa lowongan yang secara substansi berbeda. Di samping itu, penggunaan logika pencocokan pasti (*exact match*) pada atribut posisi dinilai kurang fleksibel menghadapi variasi penamaan di dunia nyata, sehingga gagal mengenali keterkaitan semantik antar-kata yang serupa. Sebagai langkah pengembangan ke depan untuk menciptakan platform rekrutmen yang lebih intuitif dan akurat, disarankan untuk mengganti pemodelan *exact match* pada teks menjadi metode berbasis kemiripan teks seperti TF-IDF, menerapkan pembobotan prioritas pada masing-masing atribut kriteria (*weighted cosine similarity*), serta memperluas ruang dimensi perbandingan data lowongan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh anggota tim penelitian Asmara Nindyah Rahati, Afifah Mutmainah, Bagus Satrio Ardianto, Fawwaz Gibran Sagita Al Firdaus, dan Miftakhu Subkhan atas dedikasi, kolaborasi, dan kontribusi penuh dalam setiap tahapan penelitian ini, mulai dari pengumpulan data, pengembangan model, hingga penulisan artikel.

REFERENSI

- [1] C. Koloman, R. Maulana, R. D. Z. Putri, and W. A. Harahap, "Sistem Rekomendasi Pekerjaan di bidang IT Menggunakan Algoritma Content-Based Filtering," *J. Creat. Student Res.*, vol. 1, no. 6, pp. 78–88, 2023.
- [2] V. M. Hersianty, E. L. Amalia, D. Puspitasari, and D. W. Wibowo, "Penerapan Algoritma Tf-Idf Dan Cosine Similarity Dalam Sistem Rekomendasi Lowongan Pekerjaan," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 1, pp. 1619–1625, 2025.
- [3] A. Rouf, Y. M. Pranoto, and E. Setyati, "Sistem Rekomendasi Pekerjaan Menggunakan Content Based Similarity," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 12, no. 2, pp. 618–626, 2023.
- [4] E. T. B. Ginting and I. Pratama, "Sistem rekomendasi jurusan SMK menggunakan metode content-based filtering di Kabupaten Sleman," *J. J. Sains Dan Teknol.*, vol. 3, no. 2, pp. 291–300, 2023.
- [5] A. Rouf, H. Asy'ari, M. Y. Urrohman, and F. D. Rahmawati, "REKOMENDASI PEKERJAAN BIDANG EKONOMI: SISTEM REKOMENDASI MENGGUNAKAN CONTENT BASED," *INTI Nusa Mandiri*, vol. 20, no. 1, pp. 56–64, 2025.
- [6] F. Novel, F. Novel, N. Bayes, and K. Kunci, "No Title," vol. 6, no. 1, pp. 65–70, 2024.
- [7] B. Jurnal, P. Sistem, M. I. Rayyan, S. Pratiwi, and S. E. Dewi, "Implementasi Sistem Temu Balik Informasi Data Cryptocurrency Berbasis Atribut Menggunakan Vector Space Model (VSM) kemampuannya dalam merepresentasikan data dalam bentuk vektor multidimensi dan sistem temu balik informasi non-teks , khususnya yang berbasis data numerik di bidang pengembang maupun peneliti lain dalam penerapan model IR berbasis atribut untuk mendukung analisis data keuangan digital serta pengambilan keputusan berbasis informasi yang lebih cerdas dan adaptif .," vol. 3, no. September, 2025.
- [8] Y. Januzaj and A. Luma, "Cosine similarity—a computing approach to match similarity between higher education programs and job market demands based on maximum number of common words," *Int. J. Emerg. Technol. Learn.*, vol. 17, no. 12, pp. 258–268, 2022.
- [9] K. Putri, N. A. Ramadlani, and L. Cahyani, "PENERAPAN ALGORITMA TF-IDF DAN COSINE SIMILARITY UNTUK QUERY PENCARIAN SOAL MATA PELAJARAN," vol. 20, no. 1, pp. 31–46, 2026.
- [10] P. Aplikasi, P. Tingkat, A. Candra, C. Lely, and H. Tri, "KEMIRIPAN ANTAR DOKUMEN TEKS MENGGUNAKAN," pp. 17–21.
- [11] I. S. Sakkinah, M. A. Fikri, and R. A. Rahma, "Sistem Rekomendasi Pekerjaan Berbasis Kompetensi Mahasiswa Menggunakan Pendekatan Content-Based Filtering," *J. Teknol. Inf. dan Multimed.*, vol. 8, no. 2, pp. 394–405, 2026.