

Pengembangan Sistem Pencocokan Curriculum Vitae Dengan Lowongan Pekerjaan Menggunakan MPNet Embedding Dan Cosine Similarity

Fitri Shafirawati^{1,*}, Rully Pramudita²

^{1,2} Fakultas Informatika dan Desain, Teknik Informatika, Universitas Bina Insani, Bekasi, Indonesia

Email: ^{1,*}fshafira0904@email.com, ²rullypramudita@binainsani.ac.id

^{*)} Email Penulis Utama

Abstrak—Proses screening Curriculum Vitae (CV) pada PT Bahagia Sukses Lestari masih dilakukan secara manual, sehingga memerlukan waktu lama, meningkatkan beban kerja HR, serta berpotensi menimbulkan bias subjektif. Berdasarkan hasil observasi, setiap pembukaan satu lowongan pekerjaan di PT Bahagia Sukses Lestari rata-rata menerima sekitar 60 CV, sedangkan proses rekrutmen hanya ditangani oleh dua staf HR. Dari total lamaran tersebut, hanya sekitar 20 lamaran yang benar-benar sesuai dengan kriteria awal pada tahap screening. Penelitian ini bertujuan untuk mengembangkan sistem screening CV berbasis web menggunakan pendekatan semantic similarity. Pengembangan sistem dilakukan menggunakan metode CRISP-DM, sedangkan proses pencocokan CV dengan lowongan pekerjaan menerapkan Natural Language Processing (NLP) menggunakan model MPNet untuk menghasilkan embedding dan cosine similarity untuk menghitung tingkat kesesuaian. Nilai similarity yang diperoleh kemudian diklasifikasikan menggunakan threshold 0,6 dengan kategori recommended untuk nilai $\geq 0,6$ dan not recommended untuk nilai $< 0,6$. Hasil pengujian dilakukan menggunakan 35 dokumen CV yang masing-masing diuji terhadap dua lowongan pekerjaan berbeda sehingga menghasilkan total 70 proses screening yang menunjukkan nilai kemiripan berkisar antara 0,26 hingga 0,81, dengan 22 CV termasuk kategori recommended dan 48 CV not recommended. Sistem mampu melakukan proses screening CV secara otomatis dengan waktu rata-rata kurang dari 1 menit untuk setiap proses seleksi. Evaluasi model menghasilkan accuracy sebesar 97%, precision 91%, recall 100%, dan F1-score 95%, yang menunjukkan kemampuan sistem dalam mengidentifikasi kandidat yang relevan. Kesimpulannya, sistem berbasis embedding MPNet dan cosine similarity terbukti mampu meningkatkan efisiensi serta objektivitas proses screening CV dan menghasilkan rekomendasi kandidat secara konsisten.

Kata Kunci: Curriculum Vitae Screening, Semantic Similarity, Natural Language Processing, Sentence Transformer MPNet, Cosine Similarity.

Abstract— The Curriculum Vitae (CV) screening process at PT Bahagia Sukses Lestari is still conducted manually, resulting in a time-consuming procedure, increased workload for Human Resources (HR) staff, and the potential for subjective bias. Based on observations, each job vacancy at PT Bahagia Sukses Lestari receives an average of 60 CV submissions, while the recruitment process is handled by only two HR staff members. Of these applications, only around 20 meet the initial screening criteria. This study aims to develop a web-based CV screening system using a semantic similarity approach. The system was developed using the CRISP-DM methodology, while the CV-job matching process applies Natural Language Processing (NLP) using the MPNet model to generate embeddings and cosine similarity to measure the degree of relevance between CVs and job vacancies. The resulting similarity scores are then classified using a threshold of 0.6, where scores ≥ 0.6 are categorized as recommended and scores < 0.6 as not recommended. The evaluation was conducted using 35 CV documents, each tested against two different job vacancies, resulting in a total of 70 screening processes. The results showed similarity scores ranging from 0.26 to 0.81, with 22 CVs classified as recommended and 48 CVs as not recommended. The system was able to perform CV screening automatically with an average processing time of less than one minute per screening process. Model evaluation achieved an accuracy of 97%, precision of 91%, recall of 100%, and an F1-score of 95%, demonstrating the system's effectiveness in identifying relevant candidates. In conclusion, the proposed system based on MPNet embeddings and cosine similarity is proven to improve the efficiency and objectivity of the CV screening process while providing consistent candidate recommendations.

Keywords: Curriculum Vitae Screening, Semantic Similarity, Natural Language Processing, Sentence Transformer MPNet, Cosine Similarity

1. PENDAHULUAN

Curriculum Vitae (CV) merupakan komponen krusial dalam proses rekrutmen karena menjadi dasar awal bagi divisi Human Resource (HR) dalam mengevaluasi kelayakan seorang kandidat. Proses rekrutmen karyawan merupakan salah satu fungsi terpenting dalam Manajemen Sumber Daya Manusia dan berperan besar dalam menentukan keberhasilan suatu perusahaan [1]. Di tengah persaingan global yang semakin ketat, menemukan kandidat yang tepat dan berkualitas menjadi faktor penentu agar perusahaan dapat tetap kompetitif dan inovatif [2]. Permasalahan muncul ketika jumlah CV yang masuk sangat besar dan harus diseleksi dalam waktu yang terbatas. Data statistik HR menunjukkan bahwa rata-rata 215-250 kandidat bersaing dalam satu lowongan, mendorong pentingnya proses seleksi yang efisien dan efektif, karena proses rekrutmen end-to-end rata-rata

berlangsung selama 43 hari [3]. Laporan *Society for Human Resource Management* mencatat bahwa rata-rata waktu perekrutan (*time to hire*) mencapai 24 hari dan meningkat hingga 44 hari pada tahun 2023 [4]. Bahkan, laporan *The Josh Bersin Company & AMS* menunjukkan bahwa pada sektor industri tertentu, waktu perekrutan dapat melebihi 67 hari [5]. Proses seleksi yang memakan waktu lama ini berisiko menyebabkan perusahaan kehilangan kandidat terbaik, di mana sekitar 29% praktisi HR melaporkan kendala tersebut akibat keterlambatan proses seleksi [4]. Titisari dan Ikhwan menyatakan bahwa proses seleksi tanpa dukungan sistem membutuhkan waktu, tenaga, dan tingkat ketelitian yang tinggi serta berpotensi menimbulkan ketidaksesuaian antara kriteria perusahaan dan hasil seleksi [6].

Hal ini sejalan dengan permasalahan yang dihadapi PT Bahagia Sukses Lestari sebagai studi kasus dalam penelitian ini. Berdasarkan hasil wawancara, dalam periode terakhir setiap pembukaan satu lowongan pekerjaan rata-rata dapat menerima sekitar 60 lamaran dari kandidat, sementara jumlah staf HR yang bertugas dalam proses rekrutmen hanya 2 orang. Dari total lamaran tersebut, hanya sekitar 20 lamaran yang benar-benar sesuai dengan kriteria awal pada tahap screening. Proses peninjauan CV masih dilakukan secara manual oleh tim HR, di mana untuk posisi tingkat managerial satu CV dapat direview secara mendalam hingga satu hari, sedangkan untuk level crew dalam satu hari rata-rata hanya sekitar lima CV yang dapat direview. Tingginya jumlah lamaran dibandingkan dengan kapasitas proses review menyebabkan proses seleksi menjadi memakan waktu, berisiko menimbulkan bias subjektif, serta berpotensi membuat kandidat yang sebenarnya memenuhi kualifikasi terlewatkan. Selain itu, variasi format dan kualitas CV yang dikirimkan kandidat juga menambah kompleksitas proses screening.

Perkembangan teknologi *Artificial Intelligence* (AI) dan *Natural Language Processing* (NLP) membuka peluang untuk membantu proses screening CV secara otomatis dan lebih efisien. NLP merupakan cabang kecerdasan buatan yang memungkinkan komputer memahami, memproses, dan menganalisis bahasa manusia dalam bentuk teks [7]. Dalam konteks rekrutmen, NLP dapat digunakan untuk mengukur tingkat kesesuaian antara isi CV pelamar dengan deskripsi lowongan pekerjaan berdasarkan makna semantik teks, bukan hanya kesamaan kata secara literal. Salah satu pendekatan NLP yang saat ini banyak digunakan adalah Sentence Transformer. Model ini mampu menghasilkan representasi vektor (*embedding*) dari teks sehingga hubungan semantik antar dokumen dapat dianalisis secara lebih kontekstual. Penelitian sebelumnya menunjukkan bahwa model *Sentence Transformer* efektif digunakan dalam berbagai tugas semantic similarity seperti *document retrieval*, *recommendation system*, dan *literature screening* [8]. Dari beberapa model *Sentence Transformer* yang tersedia, all-mpnet-base-v2 diketahui memiliki performa yang baik dalam memahami konteks kalimat dan menghasilkan representasi semantik yang lebih akurat dibandingkan model lightweight lainnya [9]. Representasi embedding yang dihasilkan oleh model all-mpnet-base-v2 kemudian dapat dibandingkan menggunakan metode cosine similarity untuk menghitung tingkat kemiripan antara CV dan lowongan pekerjaan. Cosine similarity bekerja dengan mengukur sudut antar vektor sehingga mampu menentukan tingkat kesesuaian dokumen tanpa dipengaruhi oleh panjang teks [10]. Pendekatan ini memungkinkan sistem melakukan penilaian kandidat secara lebih objektif berdasarkan kesamaan makna informasi yang terdapat pada CV dan kebutuhan pekerjaan.

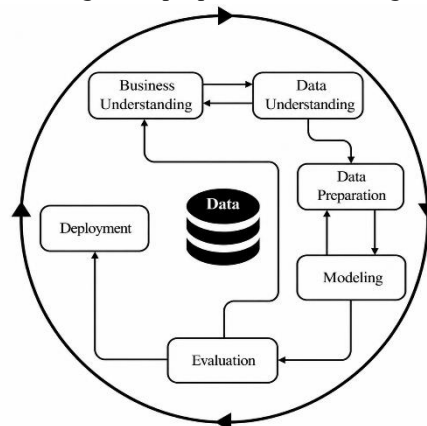
Berbagai penelitian telah menerapkan pendekatan NLP untuk mendukung proses screening CV. Affandi dan Hermanto [11] mengembangkan sistem screening CV menggunakan K-Nearest Neighbor (KNN), sedangkan Rikardo et al. [12] menerapkan pendekatan rule-based scoring berbasis kriteria HR. Penelitian lain oleh Firmansyah [13] menggunakan OCR dan pattern matching untuk mencocokkan isi CV dengan kebutuhan pekerjaan. Namun, pendekatan tersebut masih bergantung pada pencocokan kata kunci atau aturan tertentu sehingga kemampuan dalam memahami makna semantik dokumen masih terbatas. Di sisi lain, penelitian mengenai semantic similarity menunjukkan bahwa model berbasis embedding mampu memahami konteks teks secara lebih baik. Penelitian oleh Galli dan Donos serta Colangelo et al. juga menunjukkan bahwa model Sentence Transformer, khususnya MPNet, memiliki performa yang baik dalam tugas semantic similarity [8];[9]. Meskipun demikian, penelitian-penelitian tersebut masih berfokus pada domain *document retrieval*, *literature screening*, dan *recommendation system*, serta belum banyak mengkaji penerapan kombinasi *sentence embedding* MPNet dan cosine similarity pada sistem *screening* CV berbasis web. Oleh karena itu, penelitian ini mengusulkan sistem screening CV berbasis web menggunakan model all-mpnet-base-v2 dan cosine similarity untuk membantu HR dalam mengidentifikasi kandidat yang relevan berdasarkan tingkat kesesuaian semantik antara CV dan lowongan pekerjaan sehingga proses rekrutmen dapat dilakukan secara lebih optimal.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan *Natural Language Processing* untuk melakukan proses screening *Curriculum Vitae* secara otomatis berdasarkan tingkat kesesuaian terhadap lowongan pekerjaan. Model pengembangan yang digunakan dalam penelitian ini adalah *Cross Industry Standard Process for Data Mining* (CRISP-DM) karena mampu memberikan tahapan sistematis dalam pengembangan sistem berbasis data dan machine learning [14].

2.1 Tahapan Penelitian

Tahapan penelitian dilakukan berdasarkan framework CRISP-DM yang terdiri dari enam tahap utama, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment* [15].



Gambar 1. Tahapan CRISP-DM

2.1.1 Business Understanding

Tahap ini dilakukan untuk mengidentifikasi permasalahan pada proses screening CV di PT Bahagia Sukses Lestari yang masih dilakukan secara manual sehingga menimbulkan risiko berupa bias subjektif, lamanya waktu seleksi, dan meningkatnya beban kerja

2.1.2 Data Understanding

Data yang digunakan dalam penelitian ini terdiri dari 35 dokumen CV dan dua data lowongan pekerjaan. Dokumen CV memiliki variasi format, yaitu 25 CV berformat ATS dan 10 CV non-ATS, serta terdiri atas 21 CV berbahasa Indonesia dan 14 CV berbahasa Inggris. Setiap dokumen CV diuji terhadap dua lowongan pekerjaan yang berbeda sehingga menghasilkan total 70 proses pengujian untuk mengukur tingkat kesesuaian antara CV dan lowongan pekerjaan.

2.1.3 Data Preparation

Pada tahap ini dilakukan ekstraksi isi dokumen CV berformat PDF menggunakan *library* pdfplumber serta data lowongan pekerjaan yang diinput oleh pihak HR. Hasil ekstraksi kemudian melalui proses *text preprocessing* yang meliputi *cleaning* dan *case folding* untuk menghasilkan data yang lebih bersih dan konsisten sebelum diproses oleh model NLP. Proses *cleaning* dilakukan dengan menghapus karakter yang tidak relevan, seperti simbol khusus dan tanda bullet. *Text preprocessing* pada penelitian ini hanya menggunakan *cleaning* dan *case folding* karena model transformer mampu memahami konteks semantik teks melalui mekanisme *tokenization* dan *self-attention* tanpa memerlukan preprocessing linguistik yang kompleks seperti *stemming* dan *stopword removal* [16].

2.1.4 Modeling

Model *Sentence Transformer* yang telah dilatih sebelumnya (*pre-trained*) banyak digunakan untuk mengubah teks menjadi representasi vektor (*embedding*) yang mampu menangkap makna semantik suatu kalimat. Pada penelitian ini digunakan model *all-mpnet-base-v2* untuk menghasilkan *embedding* teks dari CV dan lowongan pekerjaan. Model *all-mpnet-base-v2* merupakan model transformer berbasis encoder yang mampu memahami hubungan kontekstual antar kata menggunakan mekanisme *self-attention* [17]. Setiap dokumen CV dan lowongan pekerjaan diubah menjadi representasi vektor *embedding* berdimensi 768. Representasi tersebut digunakan untuk membandingkan tingkat kesesuaian dokumen berdasarkan makna semantik, bukan hanya kesamaan kata.

Setelah *embedding* diperoleh, proses selanjutnya adalah menghitung tingkat kemiripan antara CV dan lowongan pekerjaan menggunakan cosine similarity. Metode cosine similarity menghitung sudut antara dua vektor *embedding* sehingga menghasilkan nilai *similarity* dengan rentang 0 hingga 1 [18]. Semakin tinggi nilai *similarity*, maka semakin tinggi tingkat kesesuaian kandidat terhadap lowongan pekerjaan.

$$\cos \text{sim}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} \quad (1)$$

Keterangan:

A = Vektor A, yang akan dibandingkan kemiripannya.

B = Vektor B, yang akan dibandingkan kemiripannya.

$A \cdot B$ = *Dot product* antara vektor A dan vektor B.

$\|A\|$ = Panjang vektor A.

$\| B \|$ = Panjang vektor B.

$\| A \| \| B \|$ = Hasil perkalian panjang vektor A dan panjang vektor B.

Pada penelitian ini digunakan $threshold \geq 0,6$ untuk kategori *recommended* dan $< 0,6$ untuk kategori *not recommended*. Nilai *threshold* sebesar 0,6 dipilih berdasarkan hasil eksperimen yang dilakukan pada dua nilai *threshold*, yaitu 0,5 dan 0,6. Berdasarkan hasil evaluasi, *threshold* 0,6 menghasilkan keseimbangan terbaik antara *accuracy*, *precision*, *recall*, dan *F1-score* sehingga digunakan sebagai batas klasifikasi kandidat *recommended* dan *not recommended*.

2.1.5 Evaluation

Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur performa sistem dalam mengklasifikasikan kandidat. *Ground truth* pada penelitian ini diperoleh dari hasil penilaian manual pihak HR terhadap kesesuaian CV dengan lowongan pekerjaan. Hasil klasifikasi dari sistem kemudian dibandingkan dengan hasil evaluasi manual HR untuk mengetahui tingkat akurasi sistem dalam memberikan rekomendasi kandidat.

2.1.6 Deployment

Model yang telah dikembangkan diimplementasikan ke dalam aplikasi berbasis web menggunakan framework Flask sehingga dapat digunakan oleh pelamar dan pihak HR dalam proses screening CV secara otomatis.

3. HASIL DAN PEMBAHASAN

Sistem yang dibangun dalam penelitian ini merupakan sistem screening CV berbasis web yang bertujuan untuk membantu proses seleksi awal kandidat secara otomatis. Sistem ini memanfaatkan teknik NLP untuk mengukur tingkat kesesuaian antara CV kandidat dan lowongan pekerjaan.

3.1 Ekstraksi Teks

Pada tahap ekstraksi teks, sistem hanya memproses dokumen CV dalam format .pdf berbasis teks menggunakan *library* pdfplumber. Berikut merupakan hasil ekstraksi teks dari dokumen CV.

Tabel 1. Ekstraksi Teks CV

Format CV	Isi
CV dengan format <i>Applicant Tracking System (ATS)</i> 1 halaman sebanyak 2 buah	CV Pelamar 1: Lulusan baru Teknik Mesin dari Universitas Balikpapan yang memiliki ketertarikan dan kemampuan dalam bidang Quality Control, manufaktur, dan teknik material. Saya menguasai penggunaan alat ukur presisi serta memahami prosedur inspeksi kualitas produk dan analisis hasil pengujian. Selama magang di PT. Pama Persada Nusantara pada bagian Engine Department. (...) CV Pelamar 2: Saya adalah lulusan SMK Negeri 5 Pekanbaru dengan jurusan Teknik Instalasi Tenaga Listrik. Memiliki semangat belajar yang tinggi, disiplin, dan bertanggung jawab dalam menjalankan tugas. Saya memiliki minat besar di bidang kelistrikan dan instalasi, serta siap mengembangkan keterampilan dan pengetahuan melalui pengalaman kerja langsung di lapangan. (...)
CV dengan format <i>Applicant Tracking System (ATS)</i> 2 halaman sebanyak 2 buah versi Bahasa Indonesia dan Bahasa Inggris	CV Bahasa Indonesia: Lulusan D4 Manajemen Agribisnis IPB University dengan pengalaman magang dan proyek di bidang digital marketing, business development, dan komunikasi pemasaran, baik secara remote maupun onsite. Terampil dalam riset pasar, analisis kompetitor, optimasi SEO, pengelolaan media sosial, serta pengembangan kemitraan strategis. Aktif di berbagai organisasi kampus dengan pengalaman kepemimpinan dan koordinasi tim. (...) CV Bahasa Inggris: A lifelong learner with a strong desire to acquire new skills who is deeply passionate in a Marketing Communication, Content Creator, and Customer Support Representative. Great in

Format CV	Isi
	communication, delivering impactful product presentations, and also adaptable in fast-paced environments with expertise in problem-solving and maintaining high standards of brand representation. (...)
CV dengan format Applicant Tracking System (ATS) >3 halaman versi Bahasa Inggris	7+years experience in Finance, Accounting & Tax across industries. Roles as Assistant Manager/Manager Finance with responsible reporting, taxation, audit (tax & eksternal), investor, cashflow management, financial model, forecasting and others related to finance. A Fast learner, detail, team player person, and solvable. Education TRISAKTI UNIVERSITY Jakarta, Indonesia Master Degree, Accounting (GPA 3,97) 2022 STUDY AND EVALUATION OF THE DEVELOPMENT AND RESEARCH ON THE EFFECTIVENESS OF THE AUDIT COMMITTEE DURING 2011-2020 Relevant Coursework: Good Corporate Governance, Forensic Audit, Big Data Analysis, Psychology Audit, Strategic Management. (...)
CV kreatif yang mengandung banyak gambar didalamnya sebanyak 2 buah versi Bahasa Indonesia dan Bahasa Inggris	CV Kreatif Bahasa Indonesia: Saya adalah lulusan SMAN 46 Jakarta, memiliki pengalaman kerja sebagai Kasir dan Account Executive, dengan keterampilan dalam pelayanan pelanggan, komunikasi, serta pengelolaan transaksi. Memiliki minat untuk bekerja di bidang Resepsionis atau Guest Relation Officer, dengan kemampuan dalam membangun hubungan yang baik dengan pelanggan, menangani administrasi, serta memberikan pengalaman layanan yang profesional dan ramah. (...) CV Kreatif Bahasa Inggris: I am a social media marketing strategist with 2 years work experience. Working experience in Beauty Clinic as a creative marketing, content creator, social media officer, ads specialist, live streamer and as KOL specialist. I now working in beauty SKILL clinic as Social Media Officer. I also a social media marketer who direct any brand to increase their brand value. I am a social media enthusiast, who value creativity, able to work in a team, research oriented and accustomed learning new knowledges and content creator skills. (..)

3.2 Text preprocessing

Text preprocessing dilakukan melalui tahap pembersihan dan normalisasi teks yang meliputi proses *cleaning* dan *case folding*. Proses *cleaning* dilakukan dengan menghapus karakter yang tidak relevan, seperti simbol khusus dan tanda bullet. Berikut merupakan hasil *text preprocessing* pada teks CV yang telah diekstraksi.

Tabel 2. Text preprocessing CV

Format CV	Isi
CV dengan format Applicant Tracking System (ATS) 1 halaman sebanyak 2 buah	CV Pelamar 1: lulusan baru teknik mesin dari universitas balikpapan yang memiliki ketertarikan dan kemampuan dalam bidang quality control, manufaktur, dan teknik material. saya menguasai penggunaan alat ukur presisi serta memahami prosedur inspeksi kualitas produk dan analisis hasil pengujian. selama magang di pt. pama persada nusantara pada bagian engine department. CV Pelamar 2: saya adalah lulusan smk negeri 5 pekanbaru dengan jurusan teknik instalasi tenaga listrik. memiliki semangat belajar yang tinggi, disiplin, dan bertanggung jawab dalam menjalankan tugas. saya memiliki minat besar di bidang kelistrikan dan instalasi, serta siap mengembangkan keterampilan dan pengetahuan melalui pengalaman kerja langsung di lapangan. CV Bahasa Indonesia:

Format CV	Isi
CV dengan format Applicant Tracking System (ATS) 2 halaman sebanyak 2 buah versi Bahasa Indonesia dan Bahasa Inggris	lulusan d4 manajemen agribisnis ipb university dengan pengalaman magang dan proyek di bidang digital marketing, business development, dan komunikasi pemasaran, baik secara remote maupun onsite. terampil dalam riset pasar, analisis kompetitor, optimasi seo, pengelolaan media sosial, serta pengembangan kemitraan strategis. aktif di berbagai organisasi kampus dengan pengalaman kepemimpinan dan koordinasi tim. (...) CV Bahasa Inggris: a lifelong learner with a strong desire to acquire new skills who is deeply passionate in a marketing communication, content creator, and customer support representative. great in communication, delivering impactful product presentations, and also adaptable in fast-paced environments with expertise in problem-solving and maintaining high standards of brand representation. (...)
CV dengan format Applicant Tracking System (ATS) >3 halaman versi Bahasa Inggris	7 years experience in finance,accounting tax across industries.roles as assistant manager finance with responsible reporting, taxation, audit tax eksternal , investor, cashflow management, financial model, forecasting and others related to finance. a fast learner, detail, team player person, and solvable. education trisakti university jakarta, indonesia master degree, accounting gpa 3,97 2022 study and evaluation of the development and research on the effectiveness of the audit committee during 2011-2020 relevant coursework good corporate governance, forensic audit, big data analysis, psychology audit,strategic management. (...)
CV kreatif yang mengandung banyak gambar didalamnya sebanyak 2 buah versi Bahasa Indonesia dan Bahasa Inggris	CV Kreatif Bahasa Indonesia: saya adalah lulusan sman 46 jakarta, memiliki pengalaman kerja sebagai kasir dan account executive, dengan keterampilan dalam pelayanan pelanggan, komunikasi, serta pengelolaan transaksi. memiliki minat untuk bekerja di bidang resepsionis atau guest relation officer, dengan kemampuan dalam membangun hubungan yang baik dengan pelanggan, menangani administrasi, serta memberikan pengalaman layanan yang profesional dan ramah. (...) CV Kreatif Bahasa Inggris: i am a social media marketing strategist with 2 years work experience. working experience in beauty clinic as a creative marketing, content creator, social media officer, ads specialist, live streamer and as kol specialist. i now working in beauty skill clinic as social media officer. i also a social media marketer who direct any brand to increase their brand value. i am a social media enthusiast, who value creativity, able to work in a team, research oriented and accustomed learning new knowledges and content creator skills. (...)

Berikut ini merupakan hasil dari *text preprocessing* pada isi satu lowongan pekerjaan.

Tabel 3. *Text preprocessing* Lowongan Pekerjaan

Lowongan	<i>Text preprocessing</i> Lowongan
Digital Creator Job Description Bertanggung jawab dalam merancang dan mengeksekusi strategi digital marketing yang berfokus pada pengelolaan KOL (Key Opinion Leader), influencer, serta aktivitas brand activation untuk meningkatkan awareness dan performa bisnis. Peran ini mencakup pengelolaan campaign secara end-to-end, mulai dari perencanaan, koordinasi dengan berbagai pihak, hingga analisis dan optimasi berdasarkan data performa.	bertanggung jawab dalam merancang dan mengeksekusi strategi digital marketing yang berfokus pada pengelolaan kol (key opinion leader), influencer, serta aktivitas brand activation untuk meningkatkan awareness dan performa bisnis. peran ini mencakup pengelolaan campaign secara end-to-end, mulai dari perencanaan, koordinasi dengan berbagai pihak, hingga analisis dan optimasi berdasarkan data performa. - mengembangkan dan mengeksekusi strategi digital marketing - mengelola kol dan influencer untuk mendukung campaign - menjalankan end-to-

Lowongan	Text preprocessing Lowongan
Job Responsibilities	end campaign (planning, execution, evaluation) - menyusun content plan dan mendistribusikan brief ke stakeholder terkait - mengelola hubungan dengan kol (negosiasi, scheduling, monitoring performa) - mengelola campaign digital dan aktivitas brand activation (online & offline) - mengkoordinasikan partnership, live streaming, event, dan community engagement - mengelola administrasi kampanye (agreement, invoice, report, dll.) - berkoordinasi dengan tim internal dan eksternal - menganalisis performa campaign (reach, engagement, traffic, conversion rate) - mengoptimalkan strategi berdasarkan data untuk meningkatkan efektivitas campaign
Requirements	
- Lulusan S1	
- Berpengalaman di bidang digital marketing / KOL specialist / content creator	
- Memahami strategi social media, influencer marketing, dan brand activation	
- Terbiasa mengelola end-to-end campaign digital	
- Mampu berkoordinasi dengan KOL/influencer dan stakeholder	
- Memahami analisis performa (engagement, reach, conversion rate)	
- Terbiasa menggunakan tools analitik (Google Analytics, Meta Insights, TikTok Analytics)	
- Memiliki kemampuan komunikasi dan administrasi yang baik	
- Mampu bekerja secara tim maupun individu	

3.3 Embedding all-mpnet-base-v2

Setelah dilakukan tahap *text preprocessing*, langkah selanjutnya adalah proses embedding menggunakan model all-mpnet-base-v2 untuk mengubah teks CV dan lowongan pekerjaan menjadi vektor numerik. Berikut merupakan hasil representasi teks CV dan lowongan pekerjaan yang telah melalui proses embedding.

Tabel 4. Embedding all-mpnet-base-v2

Dataset	Embedding
CV dengan format Applicant Tracking System (ATS) 1 halaman sebanyak 2 buah	CV Pelamar 1: [0.00721667],[-0.05540628],[-0.02087841], [0.05820968],[0.01219008],[...]
	CV Pelamar 2: [0.06556239],[-0.00830232],[-0.02437346], [0.05356705],[0.00224225],[...]
CV dengan format Applicant Tracking System (ATS) 2 halaman sebanyak 2 buah versi Bahasa Indonesia dan Bahasa Inggris	CV Bahasa Indonesia: [0.10733155],[-0.03309359],[-0.04111917], [-0.07213173],[0.0046944],[...]
	CV Bahasa Inggris: [0.08538854],[0.00635864],

Dataset	Embedding
	[-0.04278416], [-0.01735104], [0.04092881],[...]
CV dengan format <i>Applicant Tracking System</i> (ATS) >3 halaman versi Bahasa Inggris	[0.03013783],[0.04595632],[-0.0260141], [-0.00719874],[-0.02960218],[...]
CV kreatif yang mengandung banyak gambar didalamnya sebanyak 2 buah versi Bahasa Indonesia dan Bahasa Inggris	CV Kreatif Bahasa Indonesia: [0.10170235],[0.01441252],[-0.0330787], [0.02630804],[-0.01599985],[...] CV Kreatif Bahasa Inggris: [0.13180408],[-0.00916021],[-0.03456972] [-0.0293661],[0.03037311],[...]
Lowongan Pekerjaan Digital Creator	[0.11247922],[-0.02161372], [-0.02946564], [-0.03368217], [0.0193883],[...]

3.4 Cosine Similarity

Berikut merupakan hasil perhitungan cosine similarity antara CV dan lowongan pekerjaan setelah melalui tahap embedding.

Tabel 5. Cosine Similatrity antara CV dan Lowongan Pekerjaan

Format CV	Lowongan pekerjaan Embedding	CV Embedding	Hasil Cosine Similarity
CV dengan format <i>Applicant Tracking System</i> (ATS) 1 halaman sebanyak 2 buah	[0.11247922], [-0.02161372], [-0.02946564], [-0.03368217], [0.0193883],[...]	CV Pelamar 1: [0.00721667], [-0.05540628], [-0.02087841], [0.05820968], [0.01219008],[...]	CV Pelamar 1: Dot Product: 0.33131128549575806 Norm CV: 1.0 Norm Job: 0.9999999403953552
			Cosine Similarity: 0.33131131529808044
		CV Pelamar 2: [0.06556239], [-0.00830232], [-0.02437346], [0.05356705], [0.00224225],[...]	CV Pelamar 2: Dot Product: 0.4968847334384918 Norm CV: 0.9999998807907104 Norm Job: 0.9999999403953552
			Cosine Similarity: 0.496884822845459
CV dengan format <i>Applicant Tracking System</i> (ATS) 2 halaman sebanyak 2 buah		CV Bahasa Indonesia: [0.10733155], [-0.03309359], [-0.04111917], [-0.07213173], [0.0046944],[...]	CV Bahasa Inggris: Dot Product: 0.8118751049041748 Norm CV: 1.0 Norm Job: 1.0
			Cosine Similarity: 0.8118751049041748
		CV Bahasa Inggris: [0.08538854], [0.00635864], [-0.04278416], [-0.01735104], [0.04092881],[...]	CV Bahasa Inggris: Dot Product: 0.7109954357147217 Norm CV: 1.0 Norm Job: 0.9999999403953552
			Cosine Similarity: 0.7109954953193665
CV dengan format <i>Applicant Tracking</i>		[0.03013783], [0.04595632], [-0.0260141],	Dot Product: 0.3948865532875061

Format CV	Lowongan pekerjaan Embedding	CV Embedding	Hasil Cosine Similarity
System (ATS) >3 halaman		[-0.00719874], [-0.02960218],[...]	Norm CV: 1.0000001192092896 Norm Job: 0.9999999403953552 Cosine Similarity: 0.3948865532875061
CV kreatif yang mengandung banyak gambar didalamnya sebanyak 2 buah versi Bahasa Indonesia dan Bahasa Inggris	CV Kreatif Bahasa Indonesia:	[0.10170235], [0.01441252], [-0.0330787], [0.02630804], [-0.01599985],[...]	CV Kreatif Bahasa Indonesia: Dot Product: 0.5348684787750244 Norm CV: 1.0 Norm Job: 1.0 Cosine Similarity: 0.5348684787750244
	CV Kreatif Bahasa Inggris:	[0.13180408], [-0.00916021], [-0.03456972] [-0.0293661], [0.03037311],[...]	CV Kreatif Bahasa Inggris: Dot Product: 0.6025903820991516 Norm CV: 1.0 Norm Job: 1.0 Cosine Similarity: 0.6025903820991516

3.5 Klasifikasi Kandidat

Klasifikasi kandidat dalam penelitian ini dilakukan berdasarkan nilai threshold yang diperoleh dari hasil perhitungan cosine similarity. Kandidat dengan nilai $\geq 0,6$ dikategorikan sebagai *recommended*, yang menunjukkan tingkat kesesuaian tinggi antara CV dan deskripsi pekerjaan. Sementara itu, kandidat dengan nilai $< 0,6$ dikategorikan sebagai *not recommended*, yang menunjukkan tingkat kesesuaian rendah sehingga tidak direkomendasikan untuk posisi yang ditawarkan.

Tabel 6. Klasifikasi Kandidat

Format CV	Hasil Cosine Similarity	Klasifikasi Kategori
CV dengan format Applicant Tracking System (ATS) 1 halaman sebanyak 2 buah	CV Pelamar 1: Cosine Similarity: 0.33131131529808044	<i>Not recommended</i>
	CV Pelamar 2: Cosine Similarity: 0.496884822845459	<i>Not recommended</i>
CV dengan format Applicant Tracking System (ATS) 2 halaman sebanyak 2 buah	CV Bahasa Inggris: Cosine Similarity: 0.8118751049041748	<i>Recommended</i>
	CV Bahasa Inggris: Cosine Similarity: 0.7109954953193665	<i>Recommended</i>
CV dengan format Applicant Tracking System (ATS) >3 halaman	Cosine Similarity: 0.3948865532875061	<i>Not recommended</i>
CV kreatif yang mengandung banyak gambar didalamnya sebanyak 2 buah versi Bahasa Indonesia dan Bahasa Inggris	CV Kreatif Bahasa Indonesia: Cosine Similarity: 0.5521492958068848	<i>Not recommended</i>
	CV Kreatif Bahasa Inggris: Cosine Similarity: 0.6025903820991516	<i>Recommended</i>

3.6 Evaluasi Model all-mpnet-base-v2

Evaluasi dilakukan menggunakan dua skenario lowongan pekerjaan yang berbeda, dengan 35 dokumen CV yang diuji pada masing-masing lowongan sehingga menghasilkan total 70 proses screening. Lowongan pekerjaan yang digunakan pada kedua skenario disusun dalam bahasa Indonesia, sedangkan dokumen CV memiliki variasi bahasa, yaitu bahasa Indonesia dan bahasa Inggris. Penggunaan lebih dari satu skenario ini bertujuan untuk menghindari bias terhadap satu jenis data serta menguji kemampuan generalisasi model dalam menangani variasi karakteristik lowongan pekerjaan dan CV, termasuk perbedaan bahasa dan struktur dokumen.

Tabel 7. Evaluasi Model all-mpnet-base-v2

Lowongan	CV	Bahasa	Skor Cosine	Hasil Cosine	Hasil Klasifikasi Sistem	Ground truth (Manual)
Digital Creator	ATS	Indonesia	0,81	<i>Recommended</i>	1	1
Digital Creator	ATS	Inggris	0,71	<i>Recommended</i>	1	1
Digital Creator	ATS	Inggris	0,39	<i>Not recommended</i>	0	0
Digital Creator	Kreatif	Indonesia	0,55	<i>Not recommended</i>	0	0
Digital Creator	Kreatif	Indonesia	0,53	<i>Not recommended</i>	0	0
...	...	...	...	...	...	...
Finance	ATS	Inggris	0,80	<i>Recommended</i>	1	1
Finance	Kreatif	Indonesia	0,53	<i>Not recommended</i>	0	0
Finance	ATS	Inggris	0,63	<i>Recommended</i>	1	1
Finance	Kreatif	Indonesia	0,70	<i>Recommended</i>	1	1
Finance	ATS	Indonesia	0,67	<i>Recommended</i>	1	0

Setiap kandidat memiliki nilai skor cosine similarity yang dihasilkan oleh model all-mpnet-base-v2 dengan rentang nilai 0 hingga 1. Nilai tersebut kemudian diklasifikasikan ke dalam kategori berdasarkan threshold yang telah ditentukan. Selanjutnya, skor cosine similarity dikonversi menjadi label klasifikasi biner, yaitu 0 dan 1. Label 0 diberikan kepada kandidat yang termasuk dalam kategori *not recommended*, sedangkan label 1 diberikan kepada kandidat yang termasuk dalam kategori *recommended*. Sementara itu, Label GT Manual merepresentasikan hasil penilaian Human Resource (HR), di mana nilai 1 menunjukkan kandidat yang sesuai dengan kebutuhan lowongan dan nilai 0 menunjukkan kandidat yang tidak sesuai. Perbandingan antara hasil klasifikasi sistem dan ground truth dari HR digunakan untuk menghitung metrik evaluasi berupa *accuracy*, *precision*, *recall*, dan *F1-score*. Berikut ini tabel hasil evaluasi performa model all-mpnet-base-v2 pada kedua lowongan.

Tabel 8. Hasil Evaluasi all-mpnet-base-v2 Model

Model	Accuracy	Precision	Recall	F1 Score
all-mpnet-base-v2	0.97	0.91	1.0	0.95

Berdasarkan hasil evaluasi pada kedua lowongan, model all-mpnet-base-v2 menunjukkan performa yang sangat baik dan konsisten dengan nilai *accuracy* sebesar 0.97, *precision* 0.91, *recall* 1.00, dan *F1-score* 0.95. Nilai *recall* yang mencapai 1.00 menunjukkan bahwa seluruh kandidat yang relevan berhasil diidentifikasi, sedangkan nilai *precision* sebesar 0.91 mengindikasikan masih terdapat sedikit prediksi positif yang kurang tepat.

Pengujian juga dilakukan berdasarkan bahasa dan tipe CV. Pada CV berbahasa Indonesia, model memperoleh *accuracy* sebesar 0.95, sedangkan pada CV berbahasa Inggris mencapai 1.00. Hasil ini menunjukkan bahwa model mampu memahami konteks semantik pada kedua bahasa dengan sangat baik, meskipun performa yang sedikit lebih tinggi pada bahasa Inggris kemungkinan dipengaruhi oleh karakteristik data pelatihan model. Berdasarkan tipe CV, model memperoleh *accuracy* sebesar 0.96 pada CV ATS dan 1.00 pada CV kreatif. Temuan ini menunjukkan bahwa performa model lebih dipengaruhi oleh relevansi dan kualitas pengalaman yang tercantum dalam CV dibandingkan format penyajiannya.

Evaluasi pada masing-masing lowongan menunjukkan bahwa model mencapai *accuracy*, *precision*, *recall*, dan *F1-score* sebesar 1.00 pada lowongan Digital Creator. Sementara itu, pada lowongan Finance, model

memperoleh accuracy 0.94, precision 0.71, recall 1.00, dan F1-score 0.83. Nilai precision yang lebih rendah menunjukkan adanya false positive, yang disebabkan oleh kemiripan semantik beberapa pengalaman kerja dengan bidang Finance meskipun tidak sepenuhnya sesuai secara konteks. Secara keseluruhan, hasil tersebut menunjukkan bahwa model layak digunakan sebagai tahap awal screening CV karena mampu mengidentifikasi kandidat potensial secara efektif, meskipun validasi lanjutan masih diperlukan untuk mengurangi false positive.

3.7 Implementasi Sistem

Pada tahap ini dilakukan implementasi perangkat lunak dari sistem screening CV yang telah dibangun. Setelah proses pemodelan dan evaluasi model selesai dilakukan, langkah selanjutnya adalah mengintegrasikan model tersebut ke dalam aplikasi berbasis web yang digunakan oleh pengguna. Berikut merupakan hasil implementasi perangkat lunak dalam bentuk antarmuka aplikasi yang dibangun menggunakan Flask.

1. Halaman Submit Lamaran

Halaman ini digunakan oleh pelamar untuk mengirimkan lamaran pekerjaan. Pengguna diminta mengisi data diri serta mengunggah dokumen pendukung seperti CV sebelum mengirimkan lamaran ke sistem.

Digital Creator

Deskripsi Pekerjaan

Bertanggung jawab merancang dan mengeksekusi strategi digital marketing, termasuk pengelolaan KOL/Influencer dan brand activation, serta mengelola campaign end-to-end dari perencanaan hingga analisis performa.

Tanggung Jawab

- Mengembangkan dan mengeksekusi strategi digital marketing
- Mengelola KOL dan influencer untuk mendukung campaign
- Menjelaskan end-to-end campaign (planning, execution, evaluation)
- Menyusun content plan dan mendistribusikan brief ke stakeholder terkait
- Mengelola hubungan dengan KOL (negosiasi, scheduling, monitoring performa)
- Mengelola campaign digital dan aktivitas brand activation (online & offline)
- Mengkoordinasikan partnership, live streaming, event, dan community engagement
- Mengelola administrasi kampanye (agreement, invoice, report, dll.)
- Berkoordinasi dengan tim internal dan eksternal

Apply Sekarang

Data Diri

Nama Lengkap *

Email *

Phone Number *

Contoh : 08123456789

Tanggal Lahir *

dd/mm/yyyy

Pendidikan Terakhir

Pendidikan Terakhir *

Pilih Satu ...

Gambar 2. Halaman Submit Lamaran

2. Halaman Manajemen Lowongan

Halaman ini digunakan oleh HR untuk mengelola data lowongan pekerjaan setelah login ke dalam sistem. HR dapat menambahkan, mengubah, mencari, serta menghapus lowongan sesuai kebutuhan.

Manajemen Lowongan

Daftar Lowongan (Total: 6)

Cari judul atau deskripsi lowongan...

Semua Status

Cari

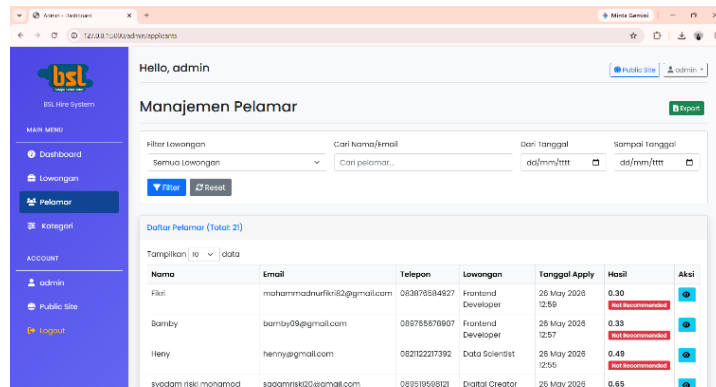
Reset

Judul Lowongan	Deskripsi	Status	Tanggal Ditawar	Jumlah Pelamar	Aksi
Backend Developer	Backend Developer untuk membangun sistem server-side yang robust dan scalable. Fokus pada performa.	Open	30 Oct 2025	1	Detail
Data Scientist	Bergabunglah dengan tim data science kami untuk membangun model machine learning yang inovatif. Anda...	Open	30 Oct 2025	1	Detail
Digital Creator	Bertanggung jawab merancang dan mengeksekusi strategi digital marketing, termasuk pengelolaan KOL/in...	Open	04 Apr 2025	1	Detail
Frontend Developer	Kami mencari Frontend Developer yang kreatif untuk membangun user interface yang menarik dan responsi...	Open	30 Oct 2025	1	Detail
Product Manager	Product Manager akan bertanggung jawab untuk mengelola product lifecycle dari konsep hingga launch...	Open	30 Oct 2025	1	Detail

Gambar 3. Halaman Manajemen Lowongan

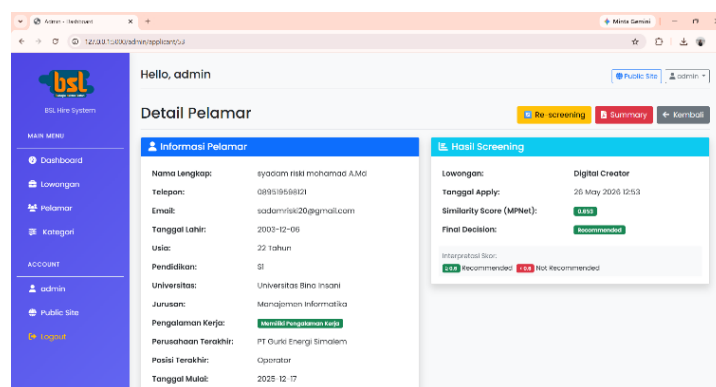
3. Halaman Manajemen Pelamar

Halaman ini digunakan oleh HR untuk mengelola data pelamar yang masuk ke dalam sistem. Pada halaman ini, HR dapat melakukan pencarian dan penyaringan data pelamar berdasarkan lowongan, nama atau email, serta rentang tanggal pendaftaran. HR juga dapat melihat detail pelamar dan dokumen CV melalui tombol aksi yang tersedia.



Gambar 5. Halaman Manajemen Pelamar

Button “Detail” pada pelamar dapat digunakan oleh HR untuk melihat informasi lengkap terkait pelamar, termasuk data diri, posisi yang dilamar, tanggal pengajuan, hasil screening, serta isi CV yang telah diunggah.



Gambar 4. Halaman Detail Pelamar

3.8 Pengujian Sistem

Pengujian perangkat lunak dilakukan untuk memastikan setiap fungsi dalam sistem berjalan sesuai dengan kebutuhan pengguna. Pada pengujian ini dilakukan dengan dua pengujian yaitu pengujian fungsional dan uji coba pengguna.

1. Pengujian Fungsional

Pengujian fungsional dilakukan pengembang untuk memastikan sistem berjalan sesuai rancangan teknis.

Tabel 9. Pengujian Fungsional

Skenario Uji	Hasil yang Diharapkan	Hasil Pengujian	Kesimpulan
HR dan pelamar mengakses web portal dengan lowongan aktif tersedia	Sistem menampilkan daftar lowongan yang berstatus aktif	Lowongan aktif ditampilkan di halaman utama	Berhasil
Pelamar mengirim lamaran dengan isi form dan unggah CV	Sistem menyimpan data dan mengalihkan halaman	Data tersimpan dan dialihkan ke halaman pemberitahuan bahwa lowongan berhasil dikirim	Berhasil
Pelamar mengunggah CV PDF yang dapat disalin	Sistem dapat membaca dan memproses CV	CV berhasil diproses	Berhasil
Pelamar mengunggah CV berupa scan/gambar	Sistem menampilkan pesan error bahwa CV tidak dapat dibaca	Muncul pesan error	Berhasil

Skenario Uji	Hasil yang Diharapkan	Hasil Pengujian	Kesimpulan
HR login dengan username dan password benar	Sistem berhasil login dan menampilkan dashboard HR	Berhasil masuk ke dashboard	Berhasil
HR menambahkan lowongan baru	Data lowongan tersimpan dan tampil di sistem	Lowongan berhasil ditambahkan	Berhasil
HR mengedit data lowongan	Perubahan data tersimpan	Data berhasil diperbarui	Berhasil
HR menghapus lowongan tanpa pelamar	Lowongan berhasil dihapus	Lowongan terhapus	Berhasil
HR menonaktifkan lowongan yang sudah memiliki pelamar	Sistem mengubah status lowongan menjadi nonaktif sehingga tidak dapat dilihat oleh pelamar	Status lowongan berhasil dinonaktifkan	Berhasil
HR melakukan filter pelamar berdasarkan lowongan	Data pelamar sesuai lowongan ditampilkan dengan ranking	Data tampil dengan ranking	Berhasil
HR mengexport data pelamar ke Excel	Sistem menghasilkan file Excel berisi data pelamar dan hasil screening	File Excel berhasil diunduh	Berhasil
HR melihat detail pelamar	Sistem menampilkan informasi lengkap pelamar dan CV	Detail tampil lengkap	Berhasil
HR mengunduh CV pelamar	Sistem mengunduh file CV asli	File CV berhasil diunduh	Berhasil
HR melihat summary pelamar	Sistem menampilkan resume screening dalam format PDF	Resume screening berhasil ditampilkan	Berhasil
HR mengubah username atau password akun admin	Sistem menyimpan perubahan akun admin	Username atau password berhasil diperbarui	Berhasil

Pengujian response time dilakukan untuk mengukur waktu yang dibutuhkan sistem dalam memproses screening CV, mulai dari pelamar menekan tombol submit hingga sistem menghasilkan nilai similarity dan kategori hasil screening. Proses yang diuji meliputi ekstraksi teks CV, *text preprocessing*, embedding menggunakan model all-mpnet-base-v2, perhitungan cosine similarity terhadap lowongan pekerjaan, serta klasifikasi kandidat berdasarkan nilai threshold yang telah ditentukan. Pengujian menggunakan beberapa CV dengan ukuran file berbeda untuk melihat pengaruh ukuran dokumen terhadap waktu pemrosesan. Hasil pengujian ditunjukkan pada tabel berikut.

Tabel 10. Pengujian Response Time

Nama CV	Ukuran File	Waktu Proses
CV 1	197 KB	53 detik
CV 2	1.25 MB	50 detik
CV 3	133 KB	20 detik
CV 4	38.8 KB	16 detik
CV 5	293 KB	41 detik

Berdasarkan hasil pengujian *response time*, waktu proses screening CV bervariasi sesuai ukuran file dan kompleksitas dokumen. Secara keseluruhan, sistem mampu menyelesaikan proses screening dalam waktu rata-rata kurang dari 1 menit per CV. Hasil pengujian menunjukkan bahwa sistem dapat menjalankan proses screening dengan baik serta menghasilkan similarity score dan kategori hasil screening secara akurat.

2. Uji Coba Pengguna

Salah satu metode pengujian yang berfokus pada keterlibatan pengguna adalah User Acceptance Testing (UAT). Sumber data dalam penelitian ini berupa data primer yang diperoleh melalui penyebaran kuesioner kepada

40 responden, dengan unit analisis berupa individu. Instrumen penelitian menggunakan skala Likert 1–5 untuk mengukur tingkat persetujuan responden terhadap sistem yang dikembangkan. Skor 1 menunjukkan “Sangat Tidak Setuju”, skor 2 “Tidak Setuju”, skor 3 “Cukup”, skor 4 “Setuju”, dan skor 5 “Sangat Setuju”.

Tabel 11. Uji Coba Pengguna

No	Kode	Pertanyaan	SS	S	C	TS	STS
1	A1	Tampilan antarmuka sistem menarik dan nyaman digunakan.	20	20	0	0	0
2	A2	Sistem mudah digunakan tanpa memerlukan pelatihan khusus.	18	22	0	0	0
3	A3	Navigasi menu dan fitur pada sistem mudah dipahami.	15	25	0	0	0
4	B1	Sistem membantu mempermudah proses rekrutmen dan screening CV.	18	22	0	0	0
5	B2	Saya merasa terbantu dengan adanya sistem ini.	15	23	1	1	0
6	B3	Saya puas dengan fungsi dan kemudahan yang diberikan oleh sistem.	15	21	4	0	0
7	C1	Fitur-fitur pada sistem berjalan sesuai dengan kebutuhan pengguna.	14	23	3	0	0
8	C2	Informasi yang ditampilkan pada sistem mudah dipahami oleh pengguna.	14	23	3	0	0
9	D1	Sistem berjalan lancar tanpa kendala atau error saat digunakan.	12	22	6	0	0
10	D2	Sistem merespon input atau proses dengan cepat.	12	22	6	0	0

Berdasarkan hasil perhitungan Skala Likert, diperoleh total skor sebesar 1728 dari skor maksimal 2000. Nilai tersebut berasal dari akumulasi jawaban 40 responden terhadap 10 pertanyaan yang diberikan pada kuesioner pengujian sistem. Setelah dilakukan perhitungan persentase menggunakan metode Skala Likert, diperoleh hasil sebesar 86,4%. Berdasarkan hasil tersebut, maka dapat disimpulkan bahwa sistem *screening* CV yang dikembangkan termasuk dalam kategori “Sangat Setuju”. Hal ini menunjukkan bahwa sistem mampu memberikan pengalaman penggunaan yang baik, mudah dipahami, membantu proses rekrutmen dan *screening* CV, serta dapat berjalan dengan lancar sesuai kebutuhan pengguna.

4. KESIMPULAN

Berdasarkan hasil penelitian dan implementasi sistem *screening Curriculum Vitae* berbasis *similarity matching* menggunakan all-mpnet-base-v2 embedding dan cosine similarity, dapat disimpulkan bahwa penelitian ini berhasil merancang dan membangun sistem *screening* CV otomatis berbasis web yang mampu membantu proses seleksi awal kandidat secara lebih efisien. Sistem yang dikembangkan dapat melakukan ekstraksi teks dari CV, menghitung tingkat kesesuaian antara CV dan lowongan pekerjaan, serta memberikan rekomendasi kandidat berdasarkan nilai similarity. Penerapan sistem ini terbukti mampu mempercepat proses seleksi secara signifikan. Jika sebelumnya proses *screening* CV pada level managerial dapat memerlukan waktu hingga satu hari untuk satu CV dan HR level crew hanya mampu meninjau sekitar lima CV per hari, maka setelah sistem diterapkan proses *screening* dapat dilakukan secara otomatis dengan waktu rata-rata kurang dari satu menit untuk setiap proses seleksi. Selain itu, penerapan metode *Natural Language Processing* menggunakan model *Sentence Transformer* all-mpnet-base-v2 dan cosine similarity terbukti mampu meningkatkan objektivitas dalam proses *screening* CV. Sistem melakukan penilaian berdasarkan kesamaan semantik antara isi CV dan deskripsi lowongan pekerjaan, sehingga evaluasi tidak hanya bergantung pada pencocokan kata secara sederhana maupun penilaian subjektif dari HR. Hasil evaluasi menunjukkan performa yang sangat baik dengan nilai *accuracy* sebesar 97%, *precision* sebesar 91%, *recall* sebesar 100%, dan *F1-score* sebesar 95%. Hasil tersebut menunjukkan bahwa sistem mampu mengidentifikasi kandidat yang relevan secara konsisten dan akurat. Selain mampu meningkatkan efisiensi dan objektivitas proses seleksi, sistem yang dikembangkan juga dapat menangani berbagai variasi format dan bahasa CV melalui proses preprocessing serta representasi semantic embedding menggunakan model all-mpnet-base-v2. Berdasarkan hasil pengujian, sistem dapat melakukan pencocokan pada CV berformat ATS maupun CV kreatif, serta pada CV berbahasa Indonesia dan Bahasa Inggris. Model MPNet mampu menghasilkan nilai *similarity* yang konsisten pada berbagai variasi tersebut, sehingga sistem dapat mendukung proses *screening* kandidat secara lebih objektif, informatif, dan sesuai dengan kebutuhan lowongan pekerjaan.

UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya penelitian ini.

DAFTAR PUSTAKA

- [1] I. M. Kharisma dan N. Wening, "Peran Rekrutmen Dan Seleksi Terhadap Kinerja Karyawan Perusahaan: Sebuah Tinjauan Literatur Sistematis," *J. E-Bis*, vol. 7, no. 1, hal. 61–80, 2023, doi: 10.37339/e-bis.v7i1.1111.
- [2] F. U. Barokah dan Ahmad Gunawan, "Strategi Rekrutmen dan Seleksi yang Efektif untuk Meningkatkan Kualitas Tenaga Kerja," *Glob. J. Lentera BITEP*, vol. 01, no. 02, hal. 60–65, 2023, [Daring]. Tersedia pada: <https://lenteranusa.id/>
- [3] L. Hafidz, "4 Data Statistik HR dalam Proses Rekrutmen," Talentics. [Daring]. Tersedia pada: <https://www.talentics.id/resources/blog/4-data-statistik-hr-dalam-proses-rekrutmen/>
- [4] Society for Human Resource Management, "SHRM benchmarking data on time-to-hire," SHRM. [Daring]. Tersedia pada: https://vendordirectory.shrm.org/company/830416/whitepapers?utm_source=chatgpt.com
- [5] The Josh Bersin Company & AMS, "Global Talent Climate: Time-to-Hire Benchmark Factbook," The Josh Bersin Company. [Daring]. Tersedia pada: https://joshbersin.com/talent-climate-series/?utm_source=chatgpt.com
- [6] M. Titisari dan K. Ikhwan, "Proses Rekrutmen dan Seleksi: Potensi Ketidakefektifan dan Faktornya," *JMK (Jurnal Manaj. dan Kewirausahaan)*, vol. 6, no. 3, hal. 11, 2021, doi: 10.32503/jmk.v6i3.1848.
- [7] M. Amien, "Sejarah dan Perkembangan Teknik Natural Language Processing (NLP) Bahasa Indonesia : Tinjauan tentang sejarah , perkembangan teknologi , dan aplikasi NLP dalam bahasa Indonesia," *ELANG J. Interdiscip. Res.*, hal. 99–105, 2023.
- [8] C. Galli dan N. Donos, "Performance of 4 Pre-Trained Sentence Transformer Models in the Semantic Query of a Systematic Review Dataset on," 2024.
- [9] M. T. Colangelo, M. Meleti, S. Guizzardi, E. Calciolari, dan C. Galli, "A comparative analysis of sentence transformer models for automated journal recommendation using PubMed metadata," *Big Data Cogn. Comput.*, vol. 9, no. 3, hal. 67, 2025.
- [10] M. Holfiana, Safrizal, A. P. Sari, dan D. Anjani, "PENERAPAN ALGORITMA COSINE SIMILARITY UNTUK MENINGKATKAN AKURASI TEMU BALIK INFORMASI PADA SISTEM Pencarian Dokumen BENcana Alam," *Glob. Res. Innov. J.*, vol. 2, no. 1, hal. 404–414, 2026.
- [11] R. M. Affandi dan D. Hermanto, "Screening CV Otomatis Dalam Natural Language Processing Menggunakan K-Nearest Neighbor Pada Aplikasi SKILLQ," *Jatekom (Jurnal Apl. Teknol. dan Komputasi)*, vol. 1, no. 1, hal. 13–21, 2025.
- [12] Rikardo, R. M. A. Shidqi, A. Sujana, S. Wibawa, U. Rohwadi, dan F. W. Handono, "SISTEM SCREENING CV OTOMATIS BERBASIS NLP DAN RULE-BASED UNTUK SELEKSI KANDIDAT BERBASIS WEB," *J. Sist. Dan Teknol. Inf.*, vol. 08, no. 1, hal. 50–64, 2026.
- [13] R. Firmansyah, "Penerapan Metode Optical Character Recognition Dan Pattern Matching Dalam Proses Screening Curriculum Vitae Rekrutmen," *STIMATA*, vol. 1, no. 1, hal. 1–12, 2025.
- [14] A. Rianti, N. W. A. Majid, dan A. Fauzi, "CRISP-DM: Metodologi Proyek Data Science," in *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB)*, Jul 2023.
- [15] F. Martinez-Plumed et al., " CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories ," *IEEE Trans. Knowl. & Data Eng.*, vol. 33, no. 08, hal. 3048–3061, Agu 2021, doi: 10.1109/TKDE.2019.2962680.
- [16] K. Ratri, R. Wardani, I. Martina, J. Fong, dan X. Wern, "Analisis Pengaruh Karakteristik Masukan Teks terhadap Kinerja MiniLM v2-L6-H384 dan BERT- Base-Uncased pada Quora Question Pairs," *J. Telemat.*, vol. 20, no. 2, hal. 138–148, 2025, doi: 10.61769/telematika.v20i2.775.
- [17] K. Song, X. Tan, T. Qin, J. Lu, dan T. Liu, "MPNet: Masked and Permuted Pre-training for Language Understanding," no. NeurIPS, hal. 1–14, 2020.
- [18] R. R. Rafi, R. A. Ramadhani, dan A. Sanjaya, "PEMANFAATAN ALGORITMA COSINE SIMILARITY UNTUK MENGKOREKSI UJIAN ESAI," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 3, hal. 2242–2254, 2025.