



OTOMATISASI IDENTIFIKASI KESENJANGAN KEAHLIAN KERJA MELALUI ANALISIS KOMPARATIF DATA LOWONGAN DAN PROFIL KANDIDAT

Iamho Pegodang Eltiuzy ¹⁾, Geovano Galan Widiatmoko Putra ²⁾, Wahyu Setiawan ³⁾,
Amalia Anjani Arifiyanti ⁴⁾

- ¹⁾ Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia
Email: 23082010133@student.upnjatim.ac.id
- ²⁾ Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia
Email: 23082010138@student.upnjatim.ac.id
- ³⁾ Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia
Email: 23082010142@student.upnjatim.ac.id
- ⁴⁾ Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia
Email: amalian_anjani.fik@upnjatim.ac.id

Abstract

This study addresses the persistent skill mismatch faced by students and new graduates in the Indonesian labor market by proposing an integrated automation framework for skill gap identification, developed as part of the TalentIQ career analytics platform. The framework combines Natural Language Processing, semantic matching, and rule-based set-comparison techniques across three datasets: 10,766 job postings, 10,720 synthetic candidate profiles, and 5,663 online courses. After text normalization and skill standardization, all job and candidate records were encoded into 384-dimensional embeddings using Sentence-BERT (all-MiniLM-L6-v2). A Two-Tower Deep Neural Network was trained on cosine-similarity-based pseudo-labels to predict candidate-job match probability, while a rule-based module compared explicit skill sets using a canonical dictionary of over 250 skills to compute matched, missing, and extra skills. Both scores were combined into a Hybrid Readiness Score (HRS). Results show that Operations & Management (36.6%), Video/Content Creator (14.2%), and Web Developer (9.1%) dominate job demand, while English, Information Architecture, and Excel are the most requested skills. The Two-Tower DNN achieved 98.60% accuracy (MAE = 0.0153) on pseudo-labeled test data. Evaluation on 500 candidate-job pairs revealed a polarized readiness distribution: 44.2% "Not Ready," 31.4% "Fairly Strong," 23.8% "Very Strong," and only 0.6% "Needs Improvement," with an average rule-based skill coverage of 21.95%. The most frequent missing skills were information architecture, social media, English, teamwork, and content creation. These findings demonstrate that combining deep semantic matching with explicit rule-based comparison produces an interpretable and actionable readiness measure, offering practical guidance for job seekers and curriculum development in Indonesia's digital creative sector.

Keywords: Skill Gap Analysis; Natural Language Processing; Semantic Matching; Two-Tower Neural Network; Labor Market Analytics.

Abstrak

Penelitian ini mengangkat persoalan skill mismatch yang dihadapi mahasiswa dan lulusan baru di pasar kerja Indonesia dengan mengusulkan kerangka otomatisasi terpadu untuk identifikasi skill gap, yang dikembangkan sebagai bagian dari platform analisis karier TalentIQ. Kerangka ini menggabungkan teknik Natural Language Processing, semantic matching, dan analisis komparatif berbasis himpunan keterampilan terhadap tiga dataset: 10.766 data lowongan kerja, 10.720 profil kandidat sintesis, dan 5.663 data kursus daring. Setelah normalisasi teks dan standarisasi keterampilan, seluruh data lowongan dan kandidat direpresentasikan sebagai vektor embedding berdimensi 384 menggunakan Sentence-BERT (all-MiniLM-L6-v2). Model Two-Tower Deep Neural Network dilatih menggunakan pseudo-label berbasis cosine similarity untuk memprediksi probabilitas kecocokan kandidat-lowongan, sementara modul rule-based membandingkan himpunan keterampilan eksplisit menggunakan kamus keterampilan kanonis (lebih dari 250 keterampilan) untuk menghasilkan keterampilan matched, missing, dan extra. Kedua skor digabungkan menjadi Hybrid Readiness Score (HRS). Hasil menunjukkan kategori Operations & Management (36,6%), Video/Content Creator (14,2%), dan Web Developer (9,1%) mendominasi kebutuhan pasar kerja, sementara English, Information Architecture, dan Excel menjadi keterampilan paling banyak dicari. Model Two-Tower DNN mencapai akurasi 98,60% (MAE = 0,0153) pada data uji. Evaluasi terhadap 500 pasangan kandidat-lowongan menunjukkan distribusi kesiapan yang polarized: 44,2% "Belum Siap", 31,4% "Cukup Kuat", 23,8% "Sangat Kuat", dan hanya 0,6% "Perlu Penguatan", dengan rata-rata rule-based skill coverage sebesar 21,95%. Keterampilan yang paling sering missing adalah information architecture, social media, english, teamwork, dan content creation. Temuan ini menunjukkan bahwa kombinasi semantic matching berbasis deep learning dengan analisis rule-based eksplisit menghasilkan ukuran kesiapan yang terukur dan dapat diinterpretasikan, sekaligus memberikan arahan praktis bagi pencari kerja dan pengembangan kurikulum pada sektor kreatif digital Indonesia.

Kata Kunci: Analisis Skill Gap; Natural Language Processing; Semantic Matching; Two-Tower Neural Network; Labor Market Analytics.



PENDAHULUAN

Revolusi industri 4.0 telah mengakselerasi transformasi digital secara masif di seluruh sektor ekonomi, termasuk pasar kerja. Menurut World Economic Forum (2020), sebanyak 85 juta pekerjaan diperkirakan akan tergantikan oleh otomatisasi pada tahun 2025, sementara 97 juta jenis pekerjaan baru akan tercipta yang menuntut kombinasi keterampilan teknis dan non-teknis yang berbeda dari sebelumnya. Pergeseran struktural ini mendorong terbentuknya lanskap kebutuhan keterampilan yang terus berubah secara dinamis, sehingga menimbulkan fenomena yang dikenal sebagai skill mismatch, yakni ketidaksesuaian antara kualifikasi yang dimiliki pencari kerja dengan kebutuhan aktual pasar kerja (McGuinness et al., 2018).

Di Indonesia, permasalahan skill mismatch menjadi isu yang semakin mendesak. Berdasarkan data (Badan Pusat Statistik, 2023), tingkat pengangguran terbuka mencapai 5,32%, dengan proporsi yang tinggi pada kelompok angkatan kerja muda berpendidikan menengah dan tinggi. Kondisi ini mencerminkan adanya ketidaksesuaian antara kualifikasi yang dimiliki lulusan dengan kebutuhan aktual pasar kerja. Permasalahan ini dapat dikategorikan ke dalam dua dimensi: (1) vertical mismatch, yaitu ketidaksesuaian antara tingkat pendidikan formal dengan kualifikasi yang dibutuhkan oleh posisi pekerjaan; dan (2) horizontal mismatch, yaitu ketidaksesuaian antara bidang studi atau keterampilan spesifik yang dikuasai kandidat dengan kompetensi yang dipersyaratkan oleh lowongan kerja (Allen, 2001). Bagi mahasiswa dan lulusan baru yang belum memiliki pengalaman profesional memadai, ketidaktahuan mengenai keterampilan spesifik yang dibutuhkan pasar kerja menjadi hambatan yang nyata dalam proses rekrutmen (Tomlinson, 2012).

Perkembangan teknik big data analytics dan Natural Language Processing (NLP) telah membuka paradigma baru dalam analisis pasar kerja. Data lowongan kerja yang tersedia secara publik di platform rekrutmen digital mengandung informasi terstruktur dan tidak terstruktur mengenai kebutuhan keterampilan, kualifikasi pendidikan, dan ekspektasi pengalaman kerja yang dapat diekstraksi dan dianalisis secara sistematis (Carnevale, Anthony P. / Jayasundera, Tanya / Repnikov, Dmitri, 2014). (Khaouja et al., 2021) dalam tinjauan sistematis terhadap 108 artikel penelitian menyimpulkan bahwa pendekatan berbasis NLP untuk identifikasi keterampilan dari iklan pekerjaan telah berkembang pesat, dengan metode berbasis deep learning menunjukkan performa terbaik untuk tugas ekstraksi keterampilan. Pendekatan berbasis data ini menawarkan keunggulan dibandingkan survei industri konvensional, di antaranya cakupan yang lebih luas dan representatif, pembaruan yang lebih cepat seiring dinamika kebutuhan industri, serta kemampuan granularitas yang lebih tinggi dalam mengidentifikasi kebutuhan keterampilan spesifik per kategori pekerjaan.

Selain ekstraksi keterampilan, teknik semantic matching berbasis representasi vektor turut memainkan peran krusial dalam analisis pasar kerja berbasis teks. Model bahasa berbasis transformer seperti BERT (Devlin et al., 2019) dan pengembangannya berupa Sentence-BERT

(Reimers & Gurevych, 2019) memungkinkan pengukuran kemiripan semantik antara profil kandidat dan deskripsi lowongan kerja secara lebih akurat dibandingkan pendekatan keyword matching konvensional, meskipun terdapat variasi penulisan keterampilan atau perbedaan terminologi. Kemampuan ini membuka peluang untuk mengotomatisasi proses pencocokan antara profil kandidat dengan lowongan kerja yang relevan, sebagai langkah awal dalam identifikasi skill gap secara otomatis.

Sejumlah penelitian sebelumnya telah mengeksplorasi berbagai aspek analisis pasar kerja berbasis NLP, seperti ekstraksi keterampilan dari deskripsi pekerjaan (Khaouja et al., 2021; Sayfullina et al., 2018), serta sistem rekomendasi pekerjaan berbasis konten yang menggabungkan NLP dengan teknik collaborative filtering (Gugnani & Misra, 2020). Namun, sebagian besar penelitian tersebut berfokus pada salah satu aspek saja, baik analisis lowongan kerja maupun sistem rekomendasi pekerjaan, tanpa mengintegrasikan keduanya ke dalam kerangka analisis skill gap yang komprehensif. Di samping itu, penelitian yang secara spesifik menyoroti konteks pasar kerja Indonesia masih sangat terbatas; sebagian besar studi menggunakan dataset dari negara maju yang tidak sepenuhnya merepresentasikan karakteristik pasar kerja Indonesia dari segi struktur industri, sistem pendidikan, dan distribusi keterampilan tenaga kerja.

Berdasarkan kondisi tersebut, penelitian ini mengusulkan sebuah kerangka analisis terpadu yang menggabungkan teknik NLP, semantic matching, dan analisis komparatif berbasis himpunan keterampilan untuk mengotomatisasi identifikasi skill gap kandidat. Penelitian ini merupakan bagian dari pengembangan sistem TalentIQ, yaitu platform analisis karier berbasis AI yang dirancang untuk membantu mahasiswa dan lulusan baru di Indonesia dalam memahami kebutuhan pasar kerja dan menyusun rencana pengembangan keterampilan yang terstruktur. Dataset yang digunakan terdiri dari ±10.000 data lowongan kerja hasil web scraping serta ±10.000 profil kandidat sintetis yang dibangun berdasarkan distribusi keterampilan aktual pada dataset lowongan, sehingga memungkinkan evaluasi metodologi skill gap identification secara sistematis dan terkontrol.

Berdasarkan uraian di atas, penelitian ini merumuskan dua pertanyaan penelitian utama. Pertama (RQ1): bagaimana profil kebutuhan keterampilan pada pasar kerja dapat diidentifikasi melalui analisis data lowongan kerja? Kedua (RQ2): bagaimana skill gap kandidat dapat diidentifikasi secara otomatis dengan membandingkan profil kandidat terhadap kebutuhan keterampilan pada lowongan yang relevan? Sejalan dengan rumusan masalah tersebut, penelitian ini bertujuan untuk: (1) memetakan profil kebutuhan keterampilan pasar kerja Indonesia secara sistematis melalui analisis komputasional terhadap dataset lowongan kerja, mencakup distribusi keterampilan dominan, kebutuhan pendidikan, dan karakteristik per kategori pekerjaan; serta (2) mengembangkan dan mengevaluasi metode otomatis untuk mengidentifikasi skill gap kandidat melalui perbandingan komparatif antara profil keterampilan kandidat dengan profil kebutuhan



keterampilan pada lowongan pekerjaan yang relevan menggunakan teknik semantic matching berbasis representasi vektor.

TINJAUAN PUSTAKA

2.1. Labor Market Analytics

Labor Market Analytics merupakan pendekatan analisis berbasis data yang digunakan untuk memahami dinamika pasar kerja, termasuk kebutuhan tenaga kerja, tren keterampilan, pola rekrutmen, dan perubahan kualifikasi pekerjaan. Dalam konteks transformasi digital, data lowongan kerja daring menjadi salah satu sumber informasi yang penting karena memuat gambaran aktual mengenai permintaan tenaga kerja dari berbagai sektor industri. Data tersebut tidak hanya berisi informasi mengenai jabatan pekerjaan, tetapi juga mencakup keterampilan teknis, kualifikasi pendidikan, pengalaman kerja, jenis pekerjaan, serta model kerja yang dibutuhkan oleh perusahaan.

Pemanfaatan data lowongan kerja daring memungkinkan proses pemetaan kebutuhan keterampilan dilakukan secara lebih cepat dibandingkan survei pasar kerja konvensional. Carnevale et al. (2014) menjelaskan bahwa data lowongan kerja daring dapat digunakan sebagai indikator pasar kerja real-time karena mampu memberikan sinyal awal mengenai perubahan kebutuhan tenaga kerja dan keterampilan. Dengan demikian, labor market analytics dapat membantu pencari kerja, lembaga pendidikan, dan pembuat kebijakan dalam menyesuaikan kompetensi tenaga kerja dengan kebutuhan industri.

2.2. Skill Gap Analysis

Skill gap atau kesenjangan keterampilan merupakan kondisi ketika keterampilan yang dimiliki individu belum sesuai dengan keterampilan yang dibutuhkan oleh pekerjaan tertentu. McGuinness et al. (2018) menjelaskan bahwa skill mismatch mencakup berbagai bentuk ketidaksesuaian dalam pasar kerja, termasuk vertical mismatch, horizontal mismatch, skill shortage, skill gap, dan skill obsolescence. Dalam konteks pencari kerja, skill gap menjadi masalah penting karena dapat menghambat kandidat dalam memenuhi kualifikasi pekerjaan yang ditawarkan oleh perusahaan.

Skill gap analysis bertujuan untuk membandingkan keterampilan yang dimiliki individu dengan keterampilan yang dibutuhkan dalam suatu pekerjaan. Hasil analisis tersebut dapat digunakan untuk menentukan keterampilan yang sudah sesuai, keterampilan yang belum dimiliki, serta keterampilan yang perlu ditingkatkan. Bagi mahasiswa dan lulusan baru, analisis ini bermanfaat untuk menyusun strategi pengembangan kompetensi secara lebih terarah, terutama dalam menghadapi persaingan pasar kerja yang semakin kompetitif.

2.3. Natural Language Processing dalam Analisis Lowongan Kerja

Natural Language Processing (NLP) merupakan bidang dalam kecerdasan buatan yang berfokus pada pemrosesan, pemahaman, dan analisis bahasa alami oleh komputer. Dalam analisis lowongan kerja, NLP berperan

penting karena sebagian besar informasi pada deskripsi pekerjaan berbentuk teks tidak terstruktur. Informasi seperti tanggung jawab pekerjaan, kualifikasi kandidat, keterampilan teknis, soft skills, dan pengalaman kerja sering ditulis dalam format naratif sehingga membutuhkan teknik pemrosesan teks agar dapat dianalisis secara sistematis.

Khaouja et al. (2021) dalam tinjauan sistematis terhadap penelitian terkait identifikasi keterampilan dari iklan pekerjaan menunjukkan bahwa metode NLP telah banyak digunakan untuk mengekstraksi keterampilan dari teks lowongan kerja. Pendekatan yang digunakan dalam penelitian terdahulu mencakup rule-based method, machine learning, deep learning, hingga pendekatan berbasis representasi bahasa. Sementara itu, Senger et al. (2024) menjelaskan bahwa skill extraction dan skill classification dari job postings menjadi salah satu topik penting dalam computational job market analysis karena dapat membantu memahami kebutuhan keterampilan secara lebih terstruktur.

2.4. Semantic Matching

Semantic matching merupakan pendekatan pencocokan berbasis makna yang bertujuan mengukur tingkat kesesuaian antara dua teks berdasarkan kedekatan semantik, bukan hanya kesamaan kata secara eksplisit. Berbeda dengan keyword matching yang hanya mencocokkan kata yang sama persis, semantic matching dapat mengenali kesamaan makna meskipun terdapat perbedaan istilah atau variasi penulisan. Misalnya, istilah "data analysis" dan "analisis data" dapat dipahami sebagai konsep yang berdekatan apabila direpresentasikan dalam ruang vektor semantik.

Perkembangan model bahasa berbasis transformer seperti BERT memberikan kontribusi besar dalam peningkatan kemampuan pemahaman konteks pada tugas NLP. Devlin et al. (2019) memperkenalkan BERT sebagai model representasi bahasa dua arah yang mampu memahami konteks kata berdasarkan sisi kiri dan kanan kalimat secara bersamaan. Selanjutnya, Reimers dan Gurevych (2019) mengembangkan Sentence-BERT, yaitu modifikasi BERT dengan arsitektur siamese network yang menghasilkan sentence embeddings sehingga dapat digunakan untuk menghitung kemiripan antarkalimat secara efisien menggunakan cosine similarity.

2.5. Job Recommendation System

Job Recommendation System merupakan sistem yang dirancang untuk memberikan rekomendasi pekerjaan kepada kandidat berdasarkan kesesuaian antara profil kandidat dan karakteristik lowongan kerja. Sistem rekomendasi pekerjaan umumnya memanfaatkan data profil pengguna, riwayat pengalaman, keterampilan, pendidikan, preferensi kerja, serta deskripsi lowongan. Pendekatan yang banyak digunakan dalam sistem rekomendasi meliputi content-based filtering, collaborative filtering, knowledge-based filtering, dan hybrid filtering.

Dalam pendekatan content-based filtering, sistem merekomendasikan lowongan berdasarkan kesamaan antara isi profil kandidat dan deskripsi pekerjaan. Pendekatan ini banyak digunakan dalam konteks rekrutmen karena



informasi mengenai kandidat dan lowongan umumnya tersedia dalam bentuk teks. Gugnani dan Misra (2020) mengembangkan pendekatan rekomendasi pekerjaan dengan memanfaatkan document embedding untuk mengekstraksi keterampilan eksplisit dan implisit dari resume serta deskripsi pekerjaan. Hasil penelitian tersebut menunjukkan bahwa representasi dokumen dapat membantu proses pencocokan kandidat dengan lowongan secara lebih efektif.

2.6. Penelitian Terkait

Sejumlah penelitian sebelumnya telah membahas pemanfaatan data lowongan kerja, Natural Language Processing (NLP), semantic matching, dan sistem rekomendasi pekerjaan dalam konteks analisis pasar kerja. Carnevale et al. (2014) menjelaskan bahwa data lowongan kerja daring dapat digunakan sebagai sumber informasi untuk mengidentifikasi tren kebutuhan keterampilan dan kualifikasi tenaga kerja secara lebih cepat dibandingkan metode survei konvensional. Sementara itu, McGuinness et al. (2018) membahas konsep skill mismatch yang mencakup ketidaksesuaian antara keterampilan tenaga kerja dengan kebutuhan aktual pasar kerja.

Dalam konteks ekstraksi keterampilan, Khaouja et al. (2021) menunjukkan bahwa pendekatan NLP, machine learning, dan deep learning telah banyak digunakan untuk mengidentifikasi keterampilan dari iklan lowongan kerja. Penelitian lain oleh Reimers dan Gurevych (2019) memperkenalkan Sentence-BERT yang memungkinkan proses pencocokan semantik antarteks secara lebih efisien melalui representasi sentence embeddings. Pendekatan ini relevan dengan penelitian ini karena proses identifikasi skill gap diawali dengan pencocokan antara profil kandidat dan lowongan kerja yang sesuai.

Selain itu, Gugnani dan Misra (2020) membahas pemanfaatan document embedding dalam sistem rekomendasi pekerjaan untuk mencocokkan resume dengan deskripsi pekerjaan. Berbeda dengan penelitian sebelumnya yang cenderung berfokus pada ekstraksi keterampilan atau rekomendasi pekerjaan secara terpisah, penelitian ini mengintegrasikan keduanya ke dalam kerangka otomatisasi identifikasi skill gap kandidat. Ringkasan penelitian terkait ditampilkan pada Tabel 1.

Tabel 1. Penelitian Terkait

Peneliti	Fokus Penelitian	Relevansi
Khaouja et al. (2021)	Skill identification dari online job ads	Dasar ekstraksi dan normalisasi skill
Reimers & Gurevych (2019)	Sentence-BERT	Dasar semantic matching profil kandidat dan lowongan
Gugnani & Misra (2020)	Job recommendation	Dasar pencocokan kandidat dan

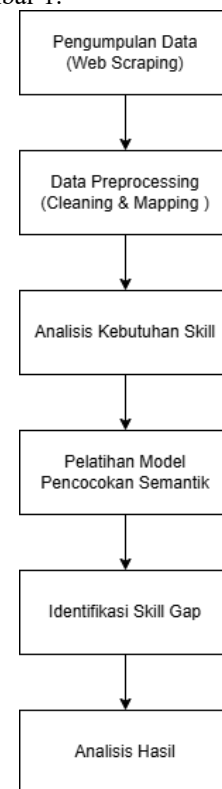
berbasis
embedding lowongan

Berdasarkan penelitian terdahulu, dapat disimpulkan bahwa data lowongan kerja daring, NLP, semantic matching, dan sistem rekomendasi pekerjaan telah banyak digunakan dalam analisis pasar kerja. Namun, sebagian besar penelitian masih berfokus pada salah satu aspek, seperti ekstraksi keterampilan atau pencocokan pekerjaan. Oleh karena itu, penelitian ini mencoba mengintegrasikan kedua aspek tersebut ke dalam kerangka otomatisasi identifikasi skill gap, yaitu dengan memetakan kebutuhan keterampilan pasar kerja, mencocokkan profil kandidat dengan lowongan kerja relevan, kemudian membandingkan keterampilan kandidat terhadap keterampilan yang dipersyaratkan oleh lowongan tersebut.

METODOLOGI PENELITIAN

3.1. Tahap Penelitian

Penelitian ini dilaksanakan melalui alur kerja yang terstruktur secara sistematis untuk memastikan validitas hasil analisis. Alur tahapan penelitian digambarkan secara tekstual pada gambar 1.



Gambar 1. Alur tahapan otomatisasi analisis kesenjangan keahlian platform TalentIQ.

Penelitian ini dilaksanakan melalui enam tahapan yang disusun secara sistematis. Tahap pertama adalah pengumpulan data, yaitu pengumpulan data lowongan kerja melalui web scraping pada platform rekrutmen daring dan pembangunan profil kandidat sintesis. Tahap kedua adalah pra-pemrosesan data, mencakup pembersihan teks, normalisasi, penghapusan duplikasi, standarisasi



keterampilan, serta komputasi representasi vektor (embedding) untuk setiap profil kandidat dan deskripsi lowongan menggunakan model Sentence-BERT all-MiniLM-L6-v2. Tahap ketiga adalah analisis kebutuhan keterampilan pasar kerja untuk menjawab RQ1, dilakukan dengan analisis frekuensi dan distribusi keterampilan pada data lowongan. Tahap keempat adalah pelatihan model pencocokan semantik, yaitu pelatihan arsitektur Two-Tower Deep Neural Network yang mempelajari pola kecocokan antara representasi vektor kandidat dan lowongan berdasarkan pseudo-label yang dihasilkan dari ambang batas cosine similarity pada data embedding. Tahap kelima adalah identifikasi skill gap, yaitu kombinasi prediksi kecocokan dari model Two-Tower DNN dengan analisis komparatif keterampilan eksplisit berbasis kamus keterampilan kanonis, untuk menjawab RQ2. Tahap keenam adalah analisis dan pembahasan hasil, mencakup interpretasi temuan dan penyusunan implikasi praktis penelitian.

3.2. Dataset

Penelitian ini menggunakan tiga dataset utama yang saling terintegrasi dalam platform TalentIQ, yaitu dataset lowongan kerja, dataset kandidat sintetis, dan dataset kursus daring. Tabel 2 menyajikan atribut inti yang digunakan secara langsung dalam analisis kebutuhan keterampilan, pencocokan kandidat-lowongan, identifikasi skill gap, dan rekomendasi pengembangan keterampilan.

Tabel 2. Skema Dataset Penelitian

Nama Dataset	Sumber / Karakteristik	Jumlah Record	Atribut Utama
Dataset Lowongan Kerja	Linkedin, Glints, Kalibr	10.766	<i>job_title, min_education, description, skills, category,</i>
Dataset Kandidat	Profil sintetis berbasis distribusi pasar kerja riil.	10.720	<i>role_category, seniority, education, skills</i>
Dataset Kursus	Coursera, edX, Dicoding	5.663	<i>course_name, category, skills_taught, level,</i>

Atribut lain, seperti identitas data, URL, informasi perusahaan, rentang gaji, dan tanggal scraping, digunakan sebagai metadata pengumpulan data sehingga tidak ditampilkan pada tabel untuk menjaga ringkasan skema dataset tetap ringkas.

3.3. Pra-Pemrosesan Data

Pra-pemrosesan data dilaksanakan dalam beberapa tahapan yang saling berkesinambungan. Pertama, dilakukan data cleaning yang mencakup penanganan missing values,

penghapusan entri duplikat, standardisasi penulisan menjadi huruf kecil (lowercase normalization), dan pembersihan karakter khusus yang tidak relevan. Tahapan ini sejalan dengan praktik standar pra-pemrosesan teks dalam analisis NLP.

Kedua, dilakukan normalisasi keterampilan untuk menangani heterogenitas representasi keterampilan dalam data teks lowongan kerja. Proses ini mencakup standardisasi berbagai variasi penulisan keterampilan yang merujuk pada konsep yang sama (misalnya, "machine learning", "ML", dan "machine-learning" distandarisasi menjadi satu bentuk baku), penghapusan redundansi keterampilan, serta pengelompokan keterampilan yang memiliki makna serupa. Proses normalisasi ini penting untuk memastikan konsistensi dalam analisis frekuensi keterampilan dan proses identifikasi skill gap selanjutnya.

Ketiga, dilakukan education mapping untuk mengelompokkan variasi penulisan tingkat pendidikan ke dalam enam kategori baku, yaitu: SD, SMP, SMA/SMK, Diploma, Sarjana S1, dan Magister S2. Keempat, dilakukan category mapping untuk mengklasifikasikan setiap lowongan kerja ke dalam kategori industri menggunakan pendekatan keyword matching berbasis daftar kata kunci yang telah ditentukan sebelumnya per kategori pekerjaan.

Untuk mendukung proses pencocokan semantik, setiap baris pada dataset lowongan kerja dan dataset kandidat dilengkapi dengan representasi vektor (embedding) berdimensi 384 yang dihasilkan menggunakan model Sentence-BERT all-MiniLM-L6-v2 (Reimers & Gurevych, 2019) dari teks gabungan atribut relevan pada masing-masing entri. Representasi vektor ini disimpan sebagai atribut tambahan (cv_embedding dan job_embedding) pada kedua dataset, sehingga proses pelatihan model pencocokan pada tahap berikutnya tidak memerlukan komputasi embedding berulang.

3.4. Analisis Kebutuhan Keterampilan Pasar Kerja (RQ1)

Untuk menjawab RQ1, analisis kebutuhan keterampilan pasar kerja dilakukan melalui empat pendekatan. Pertama, analisis frekuensi keterampilan secara keseluruhan dilakukan dengan menghitung kemunculan setiap keterampilan di seluruh dataset lowongan, kemudian diurutkan untuk mengidentifikasi keterampilan yang paling banyak diminta oleh pasar kerja.

Kedua, analisis frekuensi keterampilan per kategori pekerjaan dilakukan untuk mengidentifikasi profil keterampilan spesifik pada setiap sektor industri, sehingga dapat diketahui keterampilan yang dominan pada masing-masing kategori. Pendekatan analisis frekuensi berbasis data lowongan kerja ini relevan dengan metodologi yang digunakan oleh Carnevale et al. (2014) dalam memetakan kebutuhan keterampilan dari data iklan kerja daring.

Ketiga, analisis distribusi kebutuhan pendidikan dilakukan untuk memetakan proporsi lowongan yang mensyaratkan setiap tingkat pendidikan minimum. Keempat, analisis distribusi kategori pekerjaan dilakukan untuk menggambarkan komposisi lowongan kerja berdasarkan sektor industri dalam dataset. Keempat analisis



tersebut menghasilkan profil kebutuhan keterampilan pasar kerja yang menjadi referensi dalam tahap identifikasi skill gap pada tahap berikutnya.

3.5. Identifikasi Skill Gap Kandidat (RQ2)

Identifikasi skill gap dilakukan melalui dua komponen yang saling melengkapi, yaitu model pencocokan semantik berbasis Two-Tower Deep Neural Network dan analisis keterampilan eksplisit berbasis aturan. Kedua komponen tersebut digabungkan untuk menghasilkan Hybrid Readiness Score (HRS) sebagai ukuran tingkat kesiapan kandidat terhadap suatu lowongan kerja.

Model Two-Tower DNN menerima embedding kandidat dan lowongan berdimensi 384 sebagai masukan. Kedua representasi diproses melalui tower paralel untuk menghasilkan probabilitas kecocokan kandidat-lowongan. Karena data rekrutmen aktual tidak tersedia, pelatihan dilakukan menggunakan pseudo-label berbasis cosine similarity. Pasangan dengan cosine similarity $\geq 0,60$ diberi label Match, pasangan dengan nilai $\leq 0,30$ diberi label Not Match, sedangkan pasangan pada rentang nilai di antara keduanya tidak digunakan dalam pelatihan. Dataset pasangan dibagi menjadi data latih, validasi, dan uji dengan rasio 70:15:15. Model dilatih menggunakan binary cross-entropy dan optimizer Adam.

Komponen rule-based membandingkan keterampilan kandidat dan keterampilan yang dipersyaratkan lowongan menggunakan kamus keterampilan kanonis yang memuat lebih dari 250 keterampilan beserta variasi penulisannya. Untuk setiap pasangan kandidat-lowongan, diperoleh tiga kelompok keterampilan sebagai berikut:

$$Matched = S(c) \cap S(j)$$

$$Missing = S(j) \setminus S(c)$$

$$Extra = S(c) \setminus S(j)$$

Dengan $S(c)$ sebagai himpunan keterampilan kandidat dan $S(j)$ sebagai himpunan keterampilan yang dipersyaratkan oleh lowongan kerja. Proporsi keterampilan yang berhasil dicocokkan dihitung menggunakan rule-based skill coverage:

$$Coverage_rule = |Matched| / |S(j)|$$

Skor probabilitas model dan rule-based skill coverage kemudian digabungkan menjadi Hybrid Readiness Score:

$$HRS = (0,70 \times P_model) + (0,30 \times Coverage_rule)$$

Bobot 0,70 digunakan untuk komponen semantic matching dan bobot 0,30 digunakan untuk keterampilan eksplisit. Nilai HRS dikelompokkan menjadi empat kategori, yaitu Sangat Kuat apabila $HRS \geq 0,80$; Cukup Kuat apabila $0,65 \leq HRS < 0,80$; Perlu Penguatan apabila $0,50 \leq HRS < 0,65$; dan Belum Siap apabila $HRS < 0,50$. Keterampilan pada kelompok Missing selanjutnya

dipetakan ke dataset kursus untuk menghasilkan rekomendasi pengembangan keterampilan.

HASIL DAN PEMBAHASAN

4.1. Profil Kebutuhan Keterampilan Pasar Kerja

Analisis terhadap 10.766 data lowongan kerja menunjukkan bahwa kebutuhan tenaga kerja saat ini didominasi oleh kategori Operations & Management, yang mencakup 36,6% dari total sepuluh kategori pekerjaan teratas. Temuan ini mengindikasikan bahwa organisasi masih memiliki kebutuhan yang tinggi terhadap tenaga kerja yang mampu menjalankan fungsi operasional dan manajerial. Di sisi lain, kategori yang berkaitan dengan ekonomi digital seperti Video/Content Creator (14,2%), Web Developer (9,1%), dan Graphic Designer (7,7%) juga menunjukkan permintaan yang cukup besar, mencerminkan meningkatnya kebutuhan perusahaan terhadap transformasi digital dan pemasaran berbasis konten. Menariknya, kategori Data & AI hanya menyumbang 2,3% dari total lowongan pada sepuluh kategori teratas. Hal ini menunjukkan bahwa meskipun bidang kecerdasan buatan sedang berkembang pesat, kebutuhan tenaga kerja pada bidang tersebut masih relatif lebih kecil dibandingkan kebutuhan pada fungsi operasional, kreatif, dan pengembangan web.

Tabel 3. Sepuluh kategori pekerjaan teratas

Peringkat	Kategori	Frekuensi	Presentase %
1	Operation & Management	3942	36,6
2	Video/Content Creator	1524	14,2
3	Web Developer	981	9,1
4	Graphic Designer	824	7,7
5	Strategy & Consulting	721	6,7
6	Education & Teaching	688	6,4
7	Software Engineer	582	5,4
8	Quality Assurance (QA)	353	3,3
9	Editor & Writer	288	2,7



Peringkat	Kategori	Frekuensi	Presentase %
10	Data & AI	248	2,3

Analisis menunjukkan bahwa kombinasi keterampilan teknis dan nonteknis menjadi kebutuhan utama pasar kerja saat ini. Sepuluh keterampilan dengan frekuensi kemunculan tertinggi dalam lowongan kerja disajikan pada Tabel 4. Hasil analisis menunjukkan bahwa English (7%), Information Architecture (5,8%), Excel (4%), Teamwork (3,2%), dan Social Media (3,1%) merupakan keterampilan yang paling banyak dicari oleh perusahaan. Temuan ini mengindikasikan bahwa perusahaan tidak hanya membutuhkan kandidat yang memiliki kemampuan teknis, tetapi juga kemampuan komunikasi, kerja sama tim, serta pemahaman terhadap teknologi dan media digital. Kondisi tersebut menunjukkan bahwa daya saing tenaga kerja modern semakin ditentukan oleh kombinasi hard skills dan soft skills yang saling melengkapi.

Tabel 4. Sepuluh keterampilan pekerjaan teratas

Peringkat	Keterampilan	Frekuensi	Presentase %
1	English	6288	7
2	Information Architecture	5216	5,8
3	Excel	3646	4
4	Teamwork	2918	3,2
5	Social Media	2826	3,1
6	SEM	2187	2,4
7	communication	2148	2,4
8	NLP	1970	2,2
9	AWS	1909	2,1
10	Influencer Marketing	1789	2

4.2. Analisis Skill Gap Kandidat

Identifikasi skill gap dilakukan dengan menerapkan Hybrid Readiness Score (HRS) sebagaimana dirumuskan pada Subbab 3.5. Skor tersebut menggabungkan probabilitas kecocokan semantik dari model Two-Tower Deep Neural Network dengan rule-based skill coverage yang diperoleh melalui perbandingan keterampilan kandidat dan kebutuhan lowongan kerja.

Model Two-Tower DNN yang dilatih menggunakan pseudo-label berbasis cosine similarity mencapai akurasi

98,60% dan Mean Absolute Error (MAE) sebesar 0,0153 pada data uji. Hasil tersebut menunjukkan konsistensi model dalam mereplikasi pola kecocokan semantik yang ditetapkan melalui pseudo-label, tetapi belum dapat diartikan sebagai validasi keberhasilan rekrutmen di dunia nyata.

Evaluasi terhadap 500 pasangan kandidat-lowongan menunjukkan bahwa 56,0% pasangan diprediksi sebagai MATCH dan 44,0% sebagai NOT MATCH. Distribusi tingkat kesiapan kandidat berdasarkan HRS disajikan pada Tabel 5.

Tabel 5. Distribusi Tingkat Kesiapan Kandidat Berdasarkan Hybrid Readiness Score (n = 500)

Tingkat Kesiapan	Ambang Batas HRS	Jumlah	Presentase %
Sangat Kuat	$\geq 0,80$	119	23,8
Cukup Kuat	0,65 – 0,79	157	31,4
Perlu Penguatan	0,50 – 0,64	3	0,6
Belum Siap	$< 0,50$	221	44,2
Total		500	100

Tabel 5 menunjukkan bahwa kategori Belum Siap memiliki proporsi terbesar sebesar 44,2%, diikuti oleh Cukup Kuat sebesar 31,4% dan Sangat Kuat sebesar 23,8%. Rata-rata rule-based skill coverage sebesar 0,2195 menunjukkan bahwa hanya sekitar 22% keterampilan yang dipersyaratkan lowongan dapat dicocokkan secara eksplisit dengan profil kandidat. Temuan ini mengindikasikan adanya kesenjangan keterampilan yang cukup besar pada populasi sampel.

Secara statistik deskriptif, rata-rata jumlah keterampilan yang *matched* per pasangan kandidat-lowongan adalah 2,28 keterampilan (SD = 2,14), sementara rata-rata jumlah keterampilan yang *missing* adalah 7,58 keterampilan (SD = 5,07) dari rata-rata total keterampilan yang dipersyaratkan oleh lowongan. Nilai *rule_skill_coverage* rata-rata sebesar 0,2195 menunjukkan bahwa secara rata-rata hanya sekitar 22% keterampilan yang dipersyaratkan lowongan berhasil dicocokkan secara eksplisit dengan keterampilan pada profil kandidat, mengindikasikan kesenjangan keterampilan yang cukup signifikan pada level populasi sampel.

Keterampilan yang paling sering tergolong missing pada hasil analisis adalah information architecture (muncul pada 34,0% sampel), social media (29,2%), english (25,6%), teamwork (18,8%), dan content creation (17,8%). Dominasi keterampilan terkait pemasaran konten dan komunikasi pada daftar missing skills konsisten dengan



temuan pada Tabel 5, yang menunjukkan bahwa kategori pekerjaan *Video/Content Creator* merupakan kategori dengan jumlah sampel terbanyak (102 dari 500 pasangan) namun memiliki rata-rata *Hybrid Readiness Score* yang relatif moderat (0,528).

Tabel 6. Rata-Rata Hybrid Readiness Score per Kategori Pekerjaan (Sepuluh Kategori dengan Sampel Terbanyak)

Kategori Pekerjaan	n	Rata-rata HRS	Rata-rata Missing Skills
Video/Content Creator	102	0,528	7,57
Graphic Designer	86	0,675	6,22
Frontend Developer	61	0,705	11,21
Marketing & Growth	61	0,398	8,38
Operations & Admin	41	0,070	3,63
Software Engineer	24	0,488	10,58
Sales	17	0,017	3,29
Education & Teaching	16	0,033	1,88
Strategy & Consulting	8	0,0002	4,50
Motion Designer	8	0,576	8,88

Tabel 6 menunjukkan variasi tingkat kesiapan yang signifikan antar kategori pekerjaan. Kategori Frontend Developer menunjukkan rata-rata HRS tertinggi (0,705) di antara kategori dengan sampel signifikan, namun secara mengejutkan juga memiliki rata-rata *missing skills* yang tinggi (11,2) mengindikasikan bahwa meskipun probabilitas kecocokan semantik dari model DL tergolong tinggi, kesenjangan pada level keterampilan eksplisit tetap besar, kemungkinan karena keterampilan teknis spesifik (seperti framework atau bahasa pemrograman tertentu) lebih sulit dicocokkan secara *rule-based* dibandingkan kemiripan semantik tingkat dokumen. Sebaliknya, kategori *Operations & Admin*, *Sales*, dan *Education & Teaching* menunjukkan rata-rata HRS yang sangat rendah (< 0,10), mengindikasikan kesenjangan yang besar antara profil kandidat sintetis dan kebutuhan lowongan pada kategori-kategori tersebut dalam sampel penelitian ini.

Sebagai ilustrasi kasus individual, salah satu kandidat dengan kategori lowongan *Frontend Developer* menghasilkan probabilitas model sebesar 0,9999 (diklasifikasikan *MATCH*), dengan 6 keterampilan *matched* dan 13 keterampilan *missing* (di antaranya *angular*, *aws*, *django*, dan *flask*), menghasilkan *rule_skill_coverage* sebesar 0,316 dan HRS sebesar 0,795 (kategori Cukup Kuat). Kasus ini menggambarkan fenomena yang konsisten dengan temuan agregat pada Tabel 6: skor kecocokan semantik yang tinggi dari model DL tidak selalu berbanding lurus dengan kelengkapan keterampilan eksplisit, sehingga kombinasi kedua skor pada formulasi HRS menjadi penting untuk menghasilkan penilaian kesiapan yang lebih seimbang.

4.3. Pembahasan

Hasil penelitian menjawab RQ1 bahwa kebutuhan keterampilan pasar kerja dapat dipetakan melalui analisis frekuensi keterampilan, distribusi kategori pekerjaan, dan kebutuhan pendidikan pada data lowongan daring. Temuan menunjukkan dominasi kategori *Operations & Management* serta meningkatnya permintaan pada bidang kreatif digital dan pengembangan web. Keterampilan *English*, *Information Architecture*, dan *Excel* menjadi keterampilan yang paling sering muncul dalam data lowongan.

Terkait RQ2, kombinasi *semantic matching* berbasis *Two-Tower DNN* dan *rule-based skill matching* dapat menghasilkan penilaian kesiapan kandidat yang terukur melalui *Hybrid Readiness Score*. Pendekatan ini menggabungkan kemampuan model dalam mengenali kesamaan semantik dengan kemampuan *rule-based* dalam menunjukkan keterampilan spesifik yang masih kurang. Secara praktis, hasil penelitian dapat menjadi dasar bagi pencari kerja untuk menentukan prioritas pengembangan keterampilan dan bagi institusi pendidikan untuk menyesuaikan kurikulum dengan kebutuhan industri.

KESIMPULAN

Penelitian ini mengembangkan kerangka otomatisasi identifikasi *skill gap* pada platform *TalentIQ* melalui integrasi analisis data lowongan kerja, *semantic matching*, dan perbandingan keterampilan berbasis aturan. Kerangka tersebut memanfaatkan 10.766 data lowongan kerja, 10.720 profil kandidat sintetis, dan 5.663 data kursus daring untuk memetakan kebutuhan keterampilan pasar kerja serta mengukur kesiapan kandidat terhadap lowongan yang relevan.

Untuk RQ1, kebutuhan pasar kerja dalam dataset didominasi oleh kategori *Operations & Management* sebesar 36,6%, diikuti *Video/Content Creator* sebesar 14,2% dan *Web Developer* sebesar 9,1%. Keterampilan *English*, *Information Architecture*, dan *Excel* menjadi keterampilan yang paling banyak dicari. Untuk RQ2, kombinasi *Two-Tower DNN* dan *rule-based skill matching* menghasilkan *Hybrid Readiness Score* yang dapat menunjukkan tingkat kesiapan kandidat serta keterampilan yang masih perlu dikembangkan. Evaluasi pada 500 pasangan kandidat-lowongan menunjukkan bahwa kategori



Belum Siap memiliki proporsi terbesar sebesar 44,2%, dengan rata-rata rule-based skill coverage sebesar 21,95%.

Penelitian ini memiliki keterbatasan karena model dilatih menggunakan pseudo-label berbasis cosine similarity dan menggunakan profil kandidat sintetis, sehingga hasilnya belum merepresentasikan keberhasilan rekrutmen nyata. Penelitian selanjutnya perlu menggunakan data historis rekrutmen dan profil kandidat nyata untuk memvalidasi Hybrid Readiness Score serta mengevaluasi rekomendasi kursus yang dihasilkan.

DAFTAR PUSTAKA

- Allen, J., & van der Velden, R. (2001). Educational mismatches versus skill mismatches: Effects on wages, job satisfaction, and on-the-job search. *Oxford Economic Papers*, 53(3), 434–452. <https://doi.org/10.1093/oeq/53.3.434>
- Badan Pusat Statistik. (2023, November 6). Tingkat Pengangguran Terbuka (TPT) sebesar 5,32 persen dan rata-rata upah buruh sebesar 3,18 juta rupiah per bulan. <https://www.bps.go.id/id/pressrelease/2023/11/06/2002/tingkat-pengangguran-terbuka--tpt--sebesar-5-32-persen-dan-rata-rata-upah-buruh-sebesar-3-18-juta-rupiah-per-bulan.html>
- Carnevale, A. P., Jayasundera, T., & Repnikov, D. (2014). Understanding online job ads data: A technical report. Georgetown University Center on Education and the Workforce. <https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech.Web.pdf>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
- Gugnani, A., & Misra, H. (2020). Implicit skills extraction using document embedding and its use in job recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(08), 13286–13293. <https://doi.org/10.1609/aaai.v34i08.7038>
- Kavas, H., Serra-Vidal, M., & Wanner, L. (2025). Multilingual skill extraction for job vacancy–job seeker matching in knowledge graphs. In *Proceedings of the Generative AI and Knowledge Graph Workshop (GenAIK)* (pp. 146–155). International Committee on Computational Linguistics. <https://aclanthology.org/2025.genaik-1.15/>
- Khaouja, I., Mezzour, G., Carley, K. M., & Kassou, I. (2021). A survey on skill identification from online job ads. *IEEE Access*, 9, 118134–118153. <https://doi.org/10.1109/ACCESS.2021.3106120>
- Magron, A., Dai, A., Zhang, M., Montariol, S., & Bosselut, A. (2024). JobSkape: A framework for generating synthetic job postings to enhance skill matching. In *Proceedings of the First Workshop on Natural Language Processing for Human Resources (NLP4HR 2024)* (pp. 43–58). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.nlp4hr-1.4>
- McGuinness, S., Pouliakas, K., & Redmond, P. (2018). Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys*, 32(4), 985–1015. <https://doi.org/10.1111/joes.12254>
- Nguyen, K. C., Zhang, M., Montariol, S., & Bosselut, A. (2024). Rethinking skill extraction in the job market domain using large language models. In *Proceedings of the First Workshop on Natural Language Processing for Human Resources (NLP4HR 2024)* (pp. 27–42). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.nlp4hr-1.3>
- Qin, C., Zhu, H., Xu, T., Zhu, C., Jiang, L., Chen, E., & Xiong, H. (2018). Enhancing person-job fit for talent recruitment: An ability-aware neural network approach. In *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 25–34). <https://doi.org/10.1145/3209978.3210025>
- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 3982–3992. <https://doi.org/10.18653/v1/D19-1410>
- Sayfullina, L., Malmi, E., & Kannala, J. (2018). Learning representations for soft skill matching. *Analysis of Images, Social Networks and Texts*, 141–152. https://doi.org/10.1007/978-3-030-11027-7_13
- Senger, E., Zhang, M., van der Goot, R., & Plank, B. (2024). Deep learning-based computational job market analysis: A survey on skill extraction and classification from job postings. *Proceedings of the First Workshop on Natural Language Processing for Human Resources*, 1–16. <https://aclanthology.org/2024.nlp4hr-1.1/>
- Slokom, M. (2018). Comparing recommender systems using synthetic data. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys '18)* (pp. 548–552). ACM. <https://doi.org/10.1145/3240323.3240325>
- Tomlinson, M. (2012). Graduate employability: A review of conceptual and empirical themes. *Higher Education Policy*, 25(4), 407–431. <https://doi.org/10.1057/hep.2011.26>
- Wang, W., Wei, F., Dong, L., Bao, H., Yang, N., & Zhou, M. (2020). MiniLM: Deep self-attention distillation for task-agnostic compression of pre-



- trained transformers. arXiv.
<https://doi.org/10.48550/arXiv.2002.10957>
- World Economic Forum. (2020). The future of jobs report 2020.
https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf
- Zhang, M., Jensen, K., Sonniks, S., & Plank, B. (2022). SkillSpan: Hard and soft skill extraction from English job postings. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 4962–4984). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/2022.naacl-main.366>
- Zhao, J., Wang, J., Sigdel, M., Zhang, B., Hoang, P., Liu, M., & Korayem, M. (2021). Embedding-based recommender system for job to candidate matching on scale. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining: KDD'21 IRS Workshop.
<https://doi.org/10.48550/arXiv.2107.00221>