

Penggunaan *Jaccard Similarity Coefficient* dalam Optimasi Proses Rekrutmen Karyawan Berbasis Profil dan Kompetensi

Ivan Michael Siregar¹, Daniel Pratama², Cindy Himawan³

¹Departement of Information System, Faculty of Computer, Institut Teknologi Harapan Bangsa
Jalan Dipatiukur, Bandung, Indonesia

²Departement of Industrial Engineering, Faculty of Engineering, Institut Teknologi Harapan Bangsa
Jalan Dipatiukur, Bandung, Indonesia

³Departement of Supply Chain Management, Faculty of Engineering, Institut Teknologi Harapan Bangsa
Jalan Dipatiukur, Bandung, Indonesia

e-mail: ivan@ithb.ac.id¹, daniel.1224pratama@gmail.com², cindy_himawan@ithb.ac.id³

Received : July, 2024

Accepted : July, 2024

Published : August, 2024

Abstract

Effective employee recruitment is crucial for identifying the right employees that match the company's needs. One approach to achieving this is through algorithms that can identify potential candidates based on the required competencies. Although various methods exist to streamline recruitment, such as resume screening and job advertisement optimization, few studies focus on similarity-based methods to reduce subjectivity. This research aims to develop a Machine Learning model using the Jaccard similarity coefficient to determine potential candidates based on profile and competency similarities with existing employees. The model consists of three stages: first, identifying employees and candidates with the necessary skills; second, using the Jaccard similarity coefficient to calculate profile similarity scores and group them; third, calculating competency similarity scores and predicting shortlisted candidates. Results indicate an accuracy of 75%, precision of 71%, recall of 62%, and an F1 score of 67%, with the best stability observed in a dataset of 509 employees. An accuracy of 75% demonstrates that the model can correctly predict suitable candidates 75% of the time, sufficiently reducing subjectivity, enhancing efficiency, and assisting companies in finding the best candidates.

Keywords: competency, employee recruitment, Jaccard similarity coefficient, optimization, profile

Abstrak

Rekrutmen karyawan yang efektif sangat penting untuk menemukan karyawan yang tepat sesuai kebutuhan perusahaan. Salah satu cara untuk mencapainya adalah dengan menggunakan algoritma yang bisa mengidentifikasi kandidat potensial berdasarkan kompetensi yang dibutuhkan. Meski ada berbagai metode untuk mengefisienkan rekrutmen, seperti penyaringan resume dan optimasi iklan lowongan, belum banyak penelitian yang fokus pada metode berbasis kemiripan untuk mengurangi subjektivitas. Penelitian ini bertujuan mengembangkan model Machine Learning menggunakan Jaccard similarity coefficient untuk menentukan kandidat potensial berdasarkan kesamaan profil dan kompetensi dengan karyawan yang ada. Model ini terdiri dari tiga tahap: pertama, mengidentifikasi karyawan dan kandidat dengan keahlian sesuai kebutuhan; kedua, menggunakan Jaccard similarity coefficient untuk menghitung skor kemiripan profil dan mengelompokkan mereka; ketiga, menghitung skor kemiripan kompetensi dan memprediksi kandidat yang lolos. Hasil menunjukkan akurasi 75%, presisi 71%, recall 62%, dan f1 score 67%, dengan stabilitas terbaik pada dataset 509 karyawan. Akurasi 75% menunjukkan bahwa model

dapat memprediksi kandidat yang tepat dengan tingkat ketepatan 75%, cukup baik untuk mengurangi subjektivitas, meningkatkan efisiensi, dan membantu perusahaan menemukan kandidat terbaik.

Kata Kunci: Jaccard similarity coefficient, kompetensi, optimasi, profil, rekrutmen karyawan

1. PENDAHULUAN

Rekrutmen karyawan yang efektif dan efisien merupakan salah satu faktor penting dalam menentukan kesuksesan organisasi. Pada era digital ini, perusahaan semakin mengandalkan teknologi untuk menyaring dan memilih kandidat yang sesuai dengan kebutuhan mereka. Kandidat yang tepat tidak hanya meningkatkan produktivitas tetapi juga mengurangi biaya rekrutmen yang berulang dan masalah *turnover* karyawan [1]. Teknologi informasi khususnya *Machine Learning* dan *Artificial Intelligence*, telah membuka peluang besar untuk mengoptimalkan proses rekrutmen [2].

Penelitian terbaru menunjukkan berbagai pendekatan dalam penggunaan teknologi untuk rekrutmen. Penggunaan *Artificial Intelligence* dapat membantu penyaringan *resume* secara otomatis [3]. Penggunaan algoritma pembelajaran berbasis *support vector machine* (SVM) [4]. Penggunaan *Artificial Intelligence*, *big data*, dan otomasi memiliki dampak yang positif terhadap kinerja sumber daya manusia (SDM) [5]. perusahaan mampu mengelola SDM dan teknologi ini dengan bijak akan mendapatkan keunggulan kompetitif, meningkatkan produktivitas, dan mengoptimalkan pengelolaan SDM. Meski demikian, penelitian yang secara khusus menggunakan metode berbasis kemiripan untuk rekrutmen masih terbatas.

Penelitian ini bertujuan untuk mengembangkan model *Machine Learning* dengan menggunakan Jaccard similarity coefficient untuk menentukan kandidat potensial berdasarkan kemiripan profil dan kompetensi mereka dengan karyawan yang ada. *Jaccard similarity coefficient* merupakan metode matematis yang digunakan untuk mengukur kesamaan antara dua himpunan data, dan dalam konteks ini, digunakan untuk menghitung kemiripan antara profil kandidat dan karyawan yang ada berdasarkan atribut seperti jenis kelamin, wilayah tempat tinggal, pendidikan terakhir, lama pengalaman bekerja, gaji yang diinginkan, usia, status pernikahan, jarak tempat tinggal, dan keahlian.

Optimalisasi rekrutmen dengan menggunakan *Jaccard similarity coefficient* tidak hanya membantu dalam mengurangi waktu dan biaya rekrutmen, tetapi juga memastikan bahwa kandidat yang dipilih memiliki kecocokan yang tinggi dengan kompetensi yang dibutuhkan oleh perusahaan. Penelitian ini melibatkan tiga tahap utama, yaitu pertama memilih kandidat dan karyawan berdasarkan keterampilan yang ditentukan oleh perusahaan, kedua membuat kelompok kandidat-karyawan berdasarkan kemiripan profil, ketiga menentukan kemiripan kompetensi pada masing-masing kelompok untuk menentukan kandidat yang lolos. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam meningkatkan efektivitas dan efisiensi proses rekrutmen, mengurangi subjektivitas, serta membantu perusahaan dalam menemukan kandidat terbaik dalam memenuhi kebutuhan.

Tugas utama sebuah sistem rekomendasi adalah untuk memprediksi terhadap sebuah item yang paling cocok bagi pengguna berdasarkan preferensi dan perilaku sebelumnya. Tugas ini pada umumnya dapat dibagi menjadi dua pendekatan yaitu *content-based filtering* dan *collaborative filtering*. Pendekatan *content-based filtering* adalah berdasarkan atribut dari sebuah entitas seperti *meta-data*, deskripsi, harga, dan kategori yang selanjutnya dimanfaatkan sebagai informasi tambahan untuk memberikan informasi yang lebih tepat dan akurat [6]. Sistem rekomendasi akan memberikan hasil prediksinya sebagai alternatif usulan kepada pengguna untuk membuat keputusan. Sedangkan dengan pendekatan *collaborative filtering*, rekomendasi didasarkan kepada perilaku *user* dalam melakukan transaksi. Asumsi yang digunakan adalah bahwa *user* yang memiliki preferensi yang sama pada suatu *item* juga akan cenderung memiliki preferensi yang sama terhadap item lainnya [7], [8]. Sebagai contoh, dalam *e-commerce*, jika seorang pelanggan membeli atau melihat beberapa buku fiksi ilmiah, sistem rekomendasi akan menyarankan buku fiksi ilmiah lainnya berdasarkan genre, penulis, atau deskripsi serupa, hal ini dikategorikan sebagai *content-*

based filtering. Sedangkan, apabila seorang pelanggan yang telah membeli gadget dan headphone, maka pelanggan lainnya yang telah membeli gadget akan berpotensi besar untuk membeli headphone, hal ini dikategorikan sebagai *collaborative filtering*.

Collaborative filtering biasanya dikaitkan dengan rekomendasi produk atau konten, namun konsep yang sama juga dapat diterapkan dalam proses rekrutmen untuk menyarankan kandidat yang paling sesuai berdasarkan kesamaan dengan kandidat sebelumnya yang telah berhasil dipekerjakan atau kesamaan preferensi di antara para perekrut.

Beberapa penelitian terdahulu dalam domain tersebut telah dilakukan. Pengembangan sistem rekomendasi materi pada pembelajaran online menggunakan *collaborative filtering* [9]. Sistem ini memanfaatkan pendekatan pemrograman *hybrid* yang kompatibel dengan berbagai platform menggunakan HTML5. Penggunaan algoritma *user-based collaborative filtering* untuk memprediksi materi pembelajaran yang mungkin diminati oleh pengguna berdasarkan minat serupa dari pengguna lain [9].

Usulan sistem rekomendasi pekerjaan menggunakan *collaborative filtering* untuk membantu pencari kerja menemukan pekerjaan yang cocok [10]. Sistem ini mempertimbangkan riwayat pencarian dan aplikasi pekerjaan sebelumnya dari pencari kerja untuk memberikan rekomendasi yang sesuai dengan keterampilan Mahasiswa Teknik Informatika berdasarkan perhitungan dan analisis data keterampilan.

Integrasi *Artificial Intelligence* dalam manajemen sumber daya manusia yang telah merevolusi cara perusahaan merekrut, mengelola, dan melibatkan tenaga kerja [11]. Kualitas menjadi penting dalam penerapan *Artificial Intelligence*, sehingga dapat membantu meningkatkan produktivitas karyawan, efisiensi kerja, dan pengurangan biaya melalui otomatisasi dan analisis data. *Artificial Intelligence* dapat meningkatkan kinerja tugas rutin melalui otomatisasi, akan tetapi menghadirkan risiko yang memunculkan rasa takut dan ketidakpercayaan di kalangan perekrut [12]. Menganalisis faktor-faktor objektif yang mempengaruhi pengunduran diri karyawan serta memprediksi kemungkinan

seorang karyawan akan meninggalkan perusahaan [13]. Awal dari pengembangan model prediksi pengunduran diri karyawan yang lebih efisien, dengan potensi peningkatan melalui penggunaan *dataset* yang lebih besar dan Informasi tambahan tentang karyawan [13]. Usulan kerangka analitik yang bertujuan untuk mendukung keputusan perekrutan di dunia nyata guna meningkatkan keputusan perekrutan dan penempatan [14]. Mengidentifikasi masalah pendidikan di Indonesia yang masih rendah kualitas dan kompetensi [15]. Menjelaskan bahwa dibutuhkan analisis sejauh mana kompetensi dapat dioptimalkan untuk meningkatkan kinerja perusahaan [15]. Diperlukan pengetahuan tentang perkembangan teknologi rekomendasi yang dapat memudahkan proses pengambil keputusan serta memberikan saran yang tepat. Klasifikasi *Machine Learning* untuk prediksi penempatan karir maupun sistem rekomendasi dapat mendukung pengambilan keputusan yang baik dalam dunia pekerjaan serta pendidikan [16], [17]. Mengenali dan mengantisipasi tindakan manusia lebih awal dengan mengkorelasikan fitur masa lalu dan masa depan melalui tiga ukuran kemiripan dengan *Jaccard similarity coefficient*, yang menghasilkan akurasi tinggi dalam pengenalan dan antisipasi tindakan [18]. Mengevaluasi stabilitas kluster dalam eksperimen dengan menggunakan *Jaccard similarity coefficient* [19].

Meskipun telah banyak penelitian dilakukan untuk meningkatkan efisiensi rekrutmen, namun belum ada penelitian yang berfokus pada prediksi keahlian berdasarkan profil dan kompetensi calon karyawan. Penelitian ini memberikan kontribusi utama yaitu:

1. Membuat model sistem rekomendasi calon karyawan dengan menggunakan metode *Jaccard similarity coefficient*.
2. Menguji model dan mengevaluasi performa yang dihasilkan.
3. Menganalisis hasil rekomendasi dan mengidentifikasi kemungkinan peningkatan kinerja algoritma.

Beberapa penelitian telah dilakukan dalam penerapan teknologi untuk rekrutmen, akan tetapi masih terdapat kesenjangan yang dapat dikembangkan pada penelitian ini. Salah satu kesenjangan yaitu penelitian berbasis kemiripan yang terbatas. Penerapan *Artificial Intelligence* dan *Machine Learning* dalam rekrutmen telah

banyak dilakukan penelitian, akan tetapi pendekatan khusus yang menggunakan *Jaccard similarity coefficient* untuk mengukur kemiripan profil dan kompetensi kandidat relatif belum banyak dibahas [20]. Penelitian ini mengembangkan pendekatan baru yang berbasis kemiripan untuk meningkatkan efisiensi rekrutmen. Penelitian sebelumnya sering kali hanya mempertimbangkan beberapa atribut kandidat, sehingga hasil dari penelitian kurang mencerminkan keadaan sesungguhnya [21]. Penelitian ini mengusulkan pengukuran kemiripan yang lebih komprehensif dengan mempertimbangkan atribut seperti jenis kelamin, wilayah tempat tinggal, pendidikan terakhir, lama pengalaman bekerja, gaji yang diinginkan, usia, status pernikahan, jarak tempat tinggal, dan keahlian, yang diharapkan dapat menghasilkan seleksi kandidat yang lebih akurat. Implementasi teknologi rekrutmen berbasis teknologi diterapkan secara global terutama pada negara-negara maju [22]. Penelitian ini berfokus pada penerapan teknologi rekrutmen di Indonesia, dengan harapan dapat memberikan wawasan baru tentang bagaimana teknologi dapat diadaptasi untuk kebutuhan lokal. Penerapan rekrutmen berbasis teknologi memberikan keuntungan terhadap perusahaan [23], [24]. Penelitian ini berfokus pada penerapan *Jaccard similarity coefficient* yang berdampak terhadap optimasi rekrutmen.

Persaingan pasar tenaga kerja dan kebutuhan untuk menemukan kandidat yang tepat dengan cepat dan efisien menjadi salah satu kebutuhan dari penelitian ini. Penerapan *Jaccard similarity coefficient* dalam rekrutmen dapat membantu perusahaan mengatasi tantangan dalam proses penyaringan kandidat, sehingga dapat mengoptimalkan hasil rekrutmen. Selain itu, penelitian ini juga memberikan dasar bagi pengembangan lebih lanjut dalam penggunaan teknologi *Machine Learning* untuk rekrutmen yang lebih objektif dan *data-driven*.

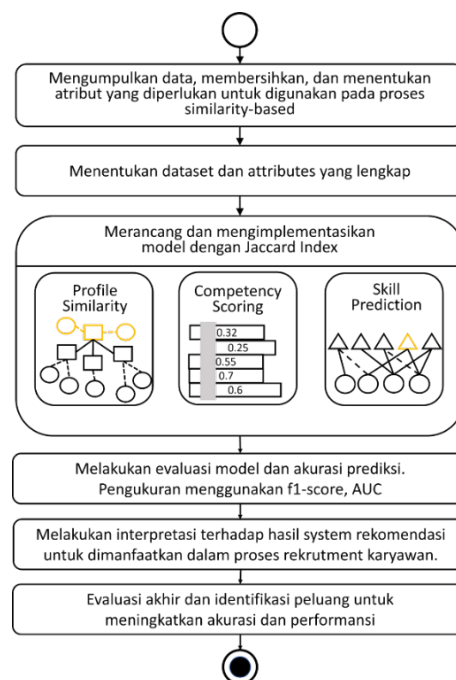
2. METODE PENELITIAN

Pada bagian ini dijelaskan metodologi penelitian yang dipakai dalam membuat rekomendasi berbasis *similarity*.

2.1 Metodologi

Penelitian dilakukan dalam tiga pembagian utama yaitu (1) persiapan *dataset* (2)

pembuatan model prediksi (3) evaluasi model dan performa. Penelitian ini dilakukan untuk mengembangkan dan mengevaluasi sistem rekomendasi berbasis *similarity* dalam proses rekrutmen karyawan. Gambar 1 menunjukkan metodologi penelitian dalam mengembangkan dan mengevaluasi sistem rekomendasi *similarity-based* dalam proses rekrutmen karyawan dengan menggunakan *Jaccard similarity coefficient* untuk *profile similarity* dan *competency scoring*. Evaluasi model menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*. Adapun metodologi penelitian lebih rinci dan terdiri dari beberapa tahap kunci sebagai berikut:



Gambar 1. Metodologi Penelitian

1. **Pengumpulan Data.** Data yang relevan dikumpulkan dari berbagai sumber, seperti profil kandidat, riwayat aplikasi pekerjaan, deskripsi pekerjaan, dan umpan balik dari proses rekrutmen sebelumnya. Data yang dikumpulkan mencakup atribut penting seperti keterampilan, pengalaman, pendidikan, dan preferensi pekerjaan.

2. **Pembersihan Data.** Data yang telah dikumpulkan kemudian dibersihkan untuk memastikan konsistensi dan kualitas. Proses pembersihan melibatkan penanganan nilai yang hilang, penghapusan data duplikat, dan standarisasi format data. Langkah ini penting

untuk menghindari bias dan kesalahan dalam model rekomendasi.

3. Penentuan Atribut. Atribut yang relevan untuk proses *similarity-based* ditentukan melalui analisis awal. Atribut ini mencakup keterampilan, kompetensi, pengalaman kerja, dan preferensi pekerjaan kandidat. Pemilihan atribut yang tepat sangat penting untuk meningkatkan akurasi sistem rekomendasi.

4. Penentuan Dataset dan Atribut Lengkap. *Dataset* yang lengkap disiapkan dengan atribut yang telah ditentukan. *Dataset* ini mencakup profil lengkap dari kandidat dan deskripsi pekerjaan yang tersedia. Penggunaan *dataset* yang lengkap dan akurat akan meningkatkan kinerja model rekomendasi.

5. Perancangan dan Implementasi Model dengan Jaccard Similarity Coefficient. Model rekomendasi dirancang dan diimplementasikan menggunakan *Jaccard similarity coefficient* untuk menghitung kesamaan antara profil kandidat dan deskripsi pekerjaan. *Jaccard similarity coefficient* digunakan karena efektivitasnya dalam mengukur kesamaan antara dua set data, yang dalam konteks ini adalah set keterampilan dan kompetensi kandidat *versus* kebutuhan pekerjaan.

6. Pengukuran Profile Similarity dan Competency Scoring. *Profile similarity* diukur dengan menghitung *Jaccard similarity coefficient* antara profil kandidat dan deskripsi pekerjaan. *Competency scoring* dilakukan untuk menilai kesesuaian kompetensi kandidat dengan kebutuhan pekerjaan. Hasil pengukuran ini digunakan untuk memberikan skor kesesuaian yang membantu dalam proses rekomendasi.

7. Prediksi Keterampilan. *Skill prediction* dilakukan untuk memprediksi keterampilan yang mungkin dimiliki kandidat berdasarkan profil mereka. Teknik ini digunakan untuk melengkapi informasi yang mungkin hilang atau tidak lengkap dalam profil kandidat.

8. Evaluasi Model dan Akurasi Prediksi. Model yang telah diimplementasikan dievaluasi untuk mengukur akurasi prediksi dan kinerjanya. Pengukuran dilakukan menggunakan metrik seperti *F1-score* dan *Area Under the Curve* (AUC). Evaluasi ini bertujuan untuk memastikan

bahwa model memberikan rekomendasi yang relevan dan akurat.

9. Interpretasi Hasil Sistem Rekomendasi. Hasil dari sistem rekomendasi diinterpretasikan untuk memahami bagaimana sistem ini dapat dimanfaatkan dalam proses rekrutmen. Interpretasi hasil membantu dalam mengidentifikasi kekuatan dan kelemahan model serta area yang perlu diperbaiki.

10. Evaluasi Akhir dan Identifikasi Peluang Peningkatan. Evaluasi akhir dilakukan untuk menilai keseluruhan kinerja dan efektivitas sistem rekomendasi. Berdasarkan hasil evaluasi, peluang untuk meningkatkan akurasi dan performansi diidentifikasi. Langkah-langkah perbaikan diusulkan untuk meningkatkan sistem rekomendasi di masa mendatang.

Metodologi penelitian ini memastikan bahwa sistem rekomendasi yang dikembangkan tidak hanya akurat dan efisien tetapi juga memberikan nilai tambah dalam proses rekrutmen karyawan.

2.2 Jaccard Similarity Coefficient

Jaccard similarity coefficient adalah pengukuran berbasis statistik untuk menentukan kemiripan atau perbedaan antara himpunan. *Jaccard similarity coefficient* dinyatakan sebagai ukuran dari hasil perbandingan jumlah elemen yang mereka miliki secara bersama-sama dengan jumlah total elemen dalam kedua himpunan tersebut [25]. Untuk dua *dataset* A dan B, maka *Jaccard similarity coefficient* adalah:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

Dimana $|A \cap B|$ adalah jumlah elemen A dan B yang beririsan, dan $|A \cup B|$ adalah jumlah elemen gabungan A dan B tanpa duplikasi. Sebagai contoh, jika diketahui himpunan $A = \{1, 2, 3, 4\}$ dan himpunan $B = \{3, 4, 5, 6\}$, maka *Jaccard similarity coefficient* $J(A, B)$ adalah:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{2}{6} = 0.333 \quad (2)$$

Dimana $|A \cap B| = \{3, 4\}$ dengan jumlah elemen pada $|A \cap B| = 2$ dan $|A \cup B| = \{1, 2, 3, 4, 5, 6\}$ dengan jumlah elemen pada $|A \cup B| = 6$. Setelah dilakukan perhitungan,

maka didapatkan hasil dari $J(A, B)$ sebesar 0.333.

Pada penelitian ini, *Jaccard similarity coefficient* digunakan sebanyak dua kali, pertama untuk menghitung kemiripan profil antara kandidat dan karyawan, berdasarkan atribut seperti jenis kelamin, tempat tinggal, tingkat pendidikan, pengalaman kerja, gaji, usia, status, jarak tinggal. Kedua, untuk menghitung kemiripan keahlian kandidat dan keahlian karyawan, dengan menggunakan sepuluh atribut keahlian seperti *critical thinking*, *data visualization*, dan *communication*.

2.3 Kebutuhan Implementasi

Dalam implementasi model seluruhnya menggunakan *framework Google Colaboratory*, menggunakan perangkat keras dan perangkat lunak untuk memastikan proses perancangan berjalan dengan lancar.

Kebutuhan dari perangkat keras sebagai berikut:

1. CPU: *High-performance virtual CPUs*
2. GPU: *NVIDIA P100s*
3. RAM: *Up to 12.72 GB*
4. Disk: *Temporary storage up to 100 GB*

Kebutuhan dari perangkat lunak sebagai berikut:

1. *Google Colaboratory*
2. *Google Spreadsheet*
3. *Python*
4. *NumPy*
5. *Pandas*
6. *Matplotlib*

3. HASIL DAN PEMBAHASAN

3.1 Dataset

Dataset yang digunakan bersumber dari sebuah perusahaan tekstil dengan jumlah karyawan mencapai 509 orang. Atribut yang melekat pada karyawan dikelompokkan menjadi dua, yaitu atribut terkait profil dan atribut terkait kompetensi, dengan perincian yang lengkap dapat dilihat pada Gambar 2.

ID	Gender	Residence Zone	Education Level	Years of Experience	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
0	E1	Other	Zone D	PhD	Advanced	CT	DV	-	-	-	-	-	-	CR
1	E2	Female	Zone I	Master's Degree	Advanced	-	-	CO	TM	TW	-	-	-	LD
2	E3	Other	Zone D	PhD	Advanced	CT	-	-	AT	-	-	-	-	CR
3	E4	Female	Zone G	High School	Beginner	-	-	CO	TM	TW	-	-	-	LD
4	E5	Female	Zone C	High School	Intermediate	CT	PS	-	CO	AT	TM	TW	-	CR
...
504	E505	Male	Zone F	PhD	Advanced	CT	-	CO	AT	-	-	-	-	LD
505	E506	Other	Zone B	High School	Beginner	-	-	CO	AT	-	-	AD	CR	LD
506	E507	Other	Zone D	Bachelor's Degree	Advanced	CT	PS	-	CO	AT	TM	TW	-	CR
507	E508	Other	Zone I	Bachelor's Degree	Intermediate	-	PS	DV	-	AT	-	TW	-	CR
508	E509	Female	Zone B	Master's Degree	Advanced	-	-	CO	-	TM	TW	-	-	CR

Gambar 2. Contoh *Dataset* Karyawan

Gambar 2 merupakan contoh *dataset* karyawan sebanyak 10 data dengan penjelasan sepuluh keterampilan (*skills*) sebagai CT = *Critical Thinking*, DV = *Data Visualization*, CO = *Communication*, AT = *Analytical Thinking*, TM = *Time Management*, TW = *Team Work*, AD = *Adaptability*, CR = *Creativity*, LD = *Leadership*.

3.2 Percobaan dan Evaluasi

Bagian ini memuat pembahasan dari data hasil penelitian yang telah disajikan. Proses pembuatan model prediksi dibagi menjadi 3 tahap yaitu (1) seleksi karyawan dan kandidat berdasarkan keterampilan (*skills*) yang dibutuhkan perusahaan (2) hitung skor kemiripan antara kandidat dan karyawan berdasarkan atribut profil (3) hitung skor kemiripan antara kandidat dan karyawan berdasarkan atribut kompetensi.

Tahap 1. Proses dimulai dengan penetapan keahlian yang dibutuhkan oleh perusahaan. Kemudian dilanjutkan dengan penyaringan karyawan yang memiliki keahlian yang sama dengan kebutuhan perusahaan, hal yang sama juga terhadap kandidat. Pada akhir tahap pertama, terseleksi karyawan dan kandidat yang memiliki keterampilan sesuai dengan kebutuhan.

Algoritma 1: *Employee and Candidate Selection*

Input:

1. *List of required skills*
2. *List of employee from data_train*
3. *List of candidate from data_test*

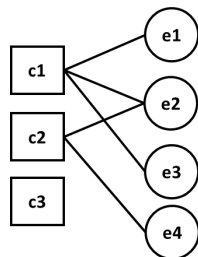
Process:

1. *Load data_train*
2. *Load data_test*
3. *Get relevant columns data_train and data_test*
4. *Determine required skills by Company*
5. *Determine the attributes used*
6. *Erase "-" from the list of skills*
7. *Select employee match required skills*
8. *Populate selected employee*
9. *Select candidate match required skills*
10. *Populate selected candidate*

Output:

1. *Selected employee of required skills*
2. *Selected candidate of required skills*

Tahap 2. Karyawan terseleksi digunakan sebagai referensi profil dalam proses pencocokan dengan profil kandidat. Profil masing-masing kandidat akan dicocokkan dengan profil karyawan terseleksi menggunakan *Jaccard similarity coefficient*. Sebagai hasilnya, untuk semua kandidat yang lolos, ditemukan peta kemiripan 1-kandidat terhadap n-karyawan. Gambar 3 menunjukkan hubungan kandidat dan karyawan. Kandidat yang memiliki hubungan dengan karyawan akan diikutsertakan, sedangkan kandidat yang tidak memiliki hubungan dengan karyawan akan dikeluarkan dari daftar kandidat.



Gambar 3. Hubungan Kandidat dan Karyawan

Pada Gambar 3. Kandidat c_1 memiliki kemiripan terhadap karyawan e_1 , e_2 , e_3 , dan kandidat c_2 memiliki kemiripan dengan e_2 , e_4 , sedangkan c_3 tidak memiliki kemiripan dengan karyawan manapun akan tidak diikutsertakan pada proses selanjutnya.

Algoritma 2: Determining Profile Similarity

Input:

1. Selected employee from step 1
2. Selected candidate from step 1

Process:

1. Load data_train
2. Load data_test
3. Select relevant columns for data_train and data_test
4. Define similarity matrix using Jaccard score
5. Calculate similarity score to find matching profiles
6. Ranking Profiles
7. Iterate data_train over each column and print unique values

Output:

1. Selected candidate having similarity with employee based on Jaccard similarity coefficient

Karyawan dengan profil yang mirip dengan kandidat akan diurutkan berdasarkan skor kemiripannya, mulai dari yang paling tinggi. Sebagai contoh, kelompok $c_1 = \{e_3, e_2, e_1\}$ artinya kandidat c_1 memiliki skor paling tinggi pada e_3 kemudian e_2 lalu e_1 . Selain itu masih ada lagi kelompok lainnya untuk $c_2 = \{e_2, e_4\}$ seperti pada Gambar 3.

Tahap 3. Proses dilakukan terhadap data kandidat dan karyawan yang sudah dikelompokkan kemiripannya. Pada masing-masing kelompok tersebut dilakukan proses kemiripan yang sama seperti tahap 2, namun kali ini berdasarkan atribut kompetensi, yaitu sepuluh keahlian. Hasil akhir dari tahap ini adalah daftar seluruh kelompok kandidat dan karyawan yang memiliki kompetensi yang mirip. Akhirnya, ditemukan data kelompok kandidat dan karyawan yang memiliki kemiripan dalam dua hal, yaitu kemiripan profil dan kemiripan kompetensi. Kandidat diprediksi memenuhi keahlian seperti yang dibutuhkan oleh perusahaan.

3.3 Training dan Testing

Dataset telah melalui proses *cleaning (labeling and encoding)* akan dilanjutkan dengan proses *training*, untuk selanjutnya siap dilatih dengan *Machine Learning*. Pembagiannya terbaik adalah dengan membagi *dataset* menjadi dua bagian yaitu data *training* sebesar 80% dan data *testing* sebesar 20%. Dengan fungsi *train_test_split* dari *library sklearn*, *dataset* dipisahkan menjadi array X dan y ke dalam 20% data *testing* (*test_size = 0.2*). Proses pemisahan *dataset* ditunjukkan pada Gambar 4.

```

from sklearn.model_selection import train_test_split

# membagi dataset menjadi training dan testing
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=1)
  
```

Gambar 4. Proses Pemisahan Dataset

3.4 Proses Implementasi Jaccard Similarity Coefficient

Satu-satunya cara untuk mengetahui apakah model yang dihasilkan sudah cukup baik adalah dengan cara menerapkannya terhadap data yang belum pernah dikenal oleh model tersebut. Untuk itu dilakukan uji coba terhadap 100 kandidat, dengan hasil uji coba adalah seperti berikut.

Tahap 1, dari 509 karyawan akan dihasilkan 55 karyawan yang memiliki keahlian sesuai kebutuhan perusahaan. Hasil seleksi karyawan dapat dilihat pada Gambar 5.

```

Employees with the required skills:
Employee ID: E6, Skills: CT, PS, Com, AT, TM, TW, Adp
Employee ID: E13, Skills: CT, PS, DV, AT, TM, Adp, Lead
Employee ID: E19, Skills: CT, PS, Com, AT, TM, TW, Adp
Employee ID: E42, Skills: CT, PS, DV, Com, AT, TM, Adp, Crvt, Lead
Employee ID: E49, Skills: CT, PS, DV, AT, TW, Adp, Lead
Employee ID: E50, Skills: CT, PS, DV, AT, TW, Adp, Lead
Employee ID: E53, Skills: CT, PS, DV, Com, AT, TM, TW, Adp, Lead
Employee ID: E78, Skills: CT, PS, Com, AT, TM, TW, Adp
Employee ID: E89, Skills: CT, PS, Com, AT, TM, TW, Adp
Employee ID: E106, Skills: CT, PS, DV, AT, TW, Adp, Lead
Employee ID: E109, Skills: CT, PS, Com, AT, TM, TW, Adp
Employee ID: E124, Skills: CT, PS, DV, AT, TW, Adp, Lead
Employee ID: E125, Skills: CT, PS, Com, AT, TM, TW, Adp
Employee ID: E128, Skills: CT, PS, DV, AT, TW, Adp, Lead
Employee ID: E132, Skills: CT, PS, DV, Com, AT, TM, Adp, Crvt, Lead
Employee ID: E168, Skills: CT, PS, DV, AT, TW, Adp, Lead
Employee ID: E184, Skills: CT, PS, Com, AT, TM, Adp, Crvt
Employee ID: E186, Skills: CT, PS, DV, AT, TW, Adp, Lead
Employee ID: E187, Skills: CT, PS, DV, Com, AT, TM, TW, Adp, Lead
Employee ID: E192, Skills: CT, PS, DV, Com, AT, TM, Adp, Crvt, Lead
Employee ID: E201, Skills: CT, PS, DV, Com, AT, TM, Adp, Crvt, Lead
Employee ID: E217, Skills: CT, PS, Com, AT, TM, Adp, Crvt
Employee ID: E229, Skills: CT, PS, Com, AT, TM, Adp, Crvt

```

Gambar 5. Hasil Seleksi Karyawan

Sama seperti karyawan, kandidat juga melalui proses seleksi berdasarkan keahlian. Pada kandidat, dari 100 kandidat dihasilkan 12 kandidat yang memiliki keahlian sesuai kebutuhan perusahaan. Hasil seleksi kandidat dapat dilihat pada Gambar 6.

```

Candidate with the required skills:
Candidate ID: C6, Skills: CT, PS, Com, AT, TM, TW, Adp
Candidate ID: C13, Skills: CT, PS, DV, AT, TW, Adp, Lead
Candidate ID: C19, Skills: CT, PS, Com, AT, TM, TW, Adp
Candidate ID: C42, Skills: CT, PS, DV, Com, AT, TM, Adp, Crvt, Lead
Candidate ID: C49, Skills: CT, PS, DV, AT, TW, Adp, Lead
Candidate ID: C50, Skills: CT, PS, DV, AT, TW, Adp, Lead
Candidate ID: C56, Skills: CT, PS, Com, AT, TM, TW, Adp
Candidate ID: C62, Skills: CT, PS, Com, AT, TM, TW, Adp
Candidate ID: C69, Skills: CT, PS, Com, AT, TM, TW, Adp
Candidate ID: C77, Skills: CT, PS, Com, AT, TM, TW, Adp
Candidate ID: C89, Skills: CT, PS, Com, AT, TM, TW, Adp
Candidate ID: C95, Skills: CT, PS, Com, AT, TM, TW, Adp

```

Gambar 6. Hasil Seleksi Kandidat

Tahap 2, dilakukan perhitungan skor kemiripan profil kandidat dan karyawan dengan menggunakan *Jaccard similarity coefficient*. Gambar 7 menunjukkan perhitungan skor kemiripan profil 12 kandidat terhadap 55 karyawan berdasarkan Informasi profil.

```

Candidate Profile for ID C6: ['C6' 'Other' 'Zone C' 'Master's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']
Top matching profiles for candidate C6 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E6' 'Other' 'Zone C' 'Master's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']
1.0 :Similarity Score , Profile: ['E181' 'Other' 'Zone C' 'Master's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']

Candidate Profile for ID C13: ['C13' 'Female' 'Zone C' 'High School' 'Intermediate' 'Low' 'Interviewing'
'Old' 'Widowed' 'Medium']
Top matching profiles for candidate C13 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E13' 'Female' 'Zone C' 'High School' 'Intermediate' 'Low' 'Interviewing'
'Old' 'Widowed' 'Medium']

Candidate Profile for ID C19: ['C19' 'Male' 'Zone D' 'Master's Degree' 'Advanced' 'High' 'Offered'
'Middle' 'Divorced' 'Medium']
Top matching profiles for candidate C19 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E19' 'Male' 'Zone D' 'Master's Degree' 'Advanced' 'High' 'Offered'
'Middle' 'Divorced' 'Medium']

Candidate Profile for ID C42: ['C42' 'Male' 'Zone G' 'Bachelor's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']
Top matching profiles for candidate C42 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E42' 'Male' 'Zone G' 'Bachelor's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']

Candidate Profile for ID C49: ['C49' 'Other' 'Zone J' 'Bachelor's Degree' 'Advanced' 'Low'
'Interviewing' 'Middle' 'Single' 'Far']
Top matching profiles for candidate C49 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E49' 'Other' 'Zone J' 'Bachelor's Degree' 'Advanced' 'Low'
'Interviewing' 'Middle' 'Single' 'Far']
1.0 :Similarity Score , Profile: ['E132' 'Other' 'Zone J' 'Bachelor's Degree' 'Advanced' 'Low'
'Interviewing' 'Middle' 'Single' 'Far']

Candidate Profile for ID C50: ['C50' 'Male' 'Zone H' 'Master's Degree' 'Beginner' 'Medium'
'Interviewing' 'Old' 'Divorced' 'Medium']
Top matching profiles for candidate C50 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E50' 'Male' 'Zone H' 'Master's Degree' 'Beginner' 'Medium'
'Interviewing' 'Old' 'Divorced' 'Medium']

```

Gambar 7. Skor Kemiripan Kandidat dan Karyawan Berdasarkan Informasi Profil

Gambar 7 terlihat kelompok kandidat dan karyawan terbentuk, contohnya $c_6 = \{e_6, e_{181}\}$, $c_{13} = \{e_{13}\}$, $c_{19} = \{c_{19}\}$, $c_{42} = \{e_{42}\}$, $c_{49} = \{e_{49}, e_{131}\}$. Kandidat yang masuk ke dalam seleksi dari tahap 2 yaitu $c_6, c_{13}, c_{19}, c_{42}, c_{49}, c_{50}, c_{56}, c_{62}, c_{69}, c_{77}, c_{89}, c_{95}$.

Tahap 3, perhitungan skor kemiripan keahlian kandidat-karyawan memberikan hasil dapat dilihat pada Gambar 8.

```

Candidate Profile for ID C6: ['C6' 'Other' 'Zone C' 'Master's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']
Top matching profiles for candidate C6 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E6' 'Other' 'Zone C' 'Master's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']
1.0 :Similarity Score , Profile: ['E181' 'Other' 'Zone C' 'Master's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']

Skill Matching for Candidate ID C6 with Employee ID E6, Candidate Skills: ['CT', 'PS', 'Com', 'AT', 'TM', 'TW', 'Adp']
1.0 :Similarity Score , Employee Skills: ['CT', 'PS', 'Com', 'AT', 'TM', 'TW', 'Adp']

Skill Matching for Candidate ID C6 with Employee ID E181, Candidate Skills: ['CT', 'PS', 'Com', 'AT', 'TM', 'TW', 'Adp']
1.0 :Similarity Score , Employee Skills: ['CT', 'PS', 'Com', 'AT', 'TM', 'TW', 'Adp']

Candidate Profile for ID C13: ['C13' 'Female' 'Zone C' 'High School' 'Intermediate' 'Low' 'Interviewing'
'Old' 'Widowed' 'Medium']
Top matching profiles for candidate C13 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E13' 'Female' 'Zone C' 'High School' 'Intermediate' 'Low' 'Interviewing'
'Old' 'Widowed' 'Medium']

Skill Matching for Candidate ID C13 with Employee ID E13, Candidate Skills: ['CT', 'PS', 'DV', 'AT', 'TM', 'TW', 'Adp', 'Lead']
1.0 :Similarity Score , Employee Skills: ['CT', 'PS', 'DV', 'AT', 'TM', 'TW', 'Adp', 'Lead']

Candidate Profile for ID C19: ['C19' 'Male' 'Zone D' 'Master's Degree' 'Advanced' 'High' 'Offered'
'Middle' 'Divorced' 'Medium']
Top matching profiles for candidate C19 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E19' 'Male' 'Zone D' 'Master's Degree' 'Advanced' 'High' 'Offered'
'Middle' 'Divorced' 'Medium']

Skill Matching for Candidate ID C19 with Employee ID E19, Candidate Skills: ['CT', 'PS', 'Com', 'AT', 'TM', 'TW', 'Adp']
1.0 :Similarity Score , Employee Skills: ['CT', 'PS', 'Com', 'AT', 'TM', 'TW', 'Adp']

Candidate Profile for ID C42: ['C42' 'Male' 'Zone G' 'Bachelor's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']
Top matching profiles for candidate C42 with a similarity score above 0.85:
1.0 :Similarity Score , Profile: ['E42' 'Male' 'Zone G' 'Bachelor's Degree' 'Advanced' 'Medium' 'Offered'
'Old' 'Widowed' 'Medium']

Skill Matching for Candidate ID C42 with Employee ID E42, Candidate Skills: ['CT', 'PS', 'DV', 'Com', 'AT', 'TM', 'TW', 'Adp', 'Crvt', 'Lead']
1.0 :Similarity Score , Employee Skills: ['CT', 'PS', 'DV', 'Com', 'AT', 'TM', 'TW', 'Adp', 'Crvt', 'Lead']

```

Gambar 8. Skor Kemiripan Kandidat dan Karyawan Berdasarkan Kompetensi

Gambar 8 terlihat kelompok kandidat dan karwan terbentuk, contohnya $c_6 = \{e_6, e_{181}\}$, $c_{13} = \{e_{13}\}$, $c_{19} = \{c_{19}\}$, $c_{42} = \{e_{42}\}$, $c_{49} = \{e_{49}, e_{131}\}$. Kandidat yang masuk ke dalam seleksi dari tahap 3 yaitu $c_{13}, c_{19}, c_{42}, c_{49}, c_{50}, c_{56}, c_{62}, c_{69}, c_{77}, c_{89}, c_{95}$. Semula terdapat c_6 , akan tetapi tidak lolos *thresholds*. Akhirnya pada tahap 3 hanya tersisa 11 kandidat dari 12 kandidat pada tahap 2.

3.5 Evaluasi Menggunakan *Confusion Matrix*

Penelitian ini menggunakan *confusion matrix* untuk mengevaluasi *Jaccard similarity coefficient*. Dengan menggunakan TP, FP, FN, dan TN dari *confusion matrix*, kita dapat dengan mudah menghitung *Jaccard similarity coefficient* untuk menilai seberapa baik model mengidentifikasi kelas positif. TP = *True Positive*, FP = *False Positive*, FN = *False Negative*, TN = *True Negative*.

Accuracy, adalah proporsi dari prediksi yang benar (baik prediksi positif maupun negatif) terhadap total prediksi yang dilakukan. Ini adalah ukuran seberapa sering model membuat prediksi yang benar secara keseluruhan.

Precision, adalah proporsi dari prediksi positif yang benar terhadap semua prediksi positif yang

dibuat oleh model. Ini mengukur seberapa banyak dari prediksi positif model yang benar-benar positif.

Recall, adalah proporsi dari kasus positif yang sebenarnya yang berhasil diidentifikasi oleh model sebagai positif. Ini mengukur kemampuan model untuk menemukan semua kasus positif.

F1 Score, adalah rata-rata harmonis dari *Precision* dan *Recall*. Ini memberikan ukuran keseimbangan antara *Precision* dan *Recall*. *F1 Score* berguna ketika kita membutuhkan keseimbangan antara keduanya dan tidak ingin hanya mengoptimalkan satu metrik dengan mengorbankan yang lain.

Tabel 1. Hasil Prediksi Kandidat pada *Confusion Matrix*

	Predicted (+)	Predicted (-)
Actual (+)	50 (TP)	30 (FN)
Actual (-)	20 (FP)	100 (TN)

Tabel 1 menampilkan hasil dari prediksi kandidat dengan menggunakan *confusion matrix*.

```

1 # Calculate confusion metrics
2 accuracy = (TP + TN) / (TP + FP + FN + TN) if (TP + FP + FN + TN) > 0 else 0.0
3 precision = TP / (TP + FP) if (TP + FP) > 0 else 0.0
4 recall = TP / (TP + FN) if (TP + FN) > 0 else 0.0
5 f1_score = 2 * (precision * recall) / (precision + recall) if (precision + recall) > 0 else 0.0
6
7 # Print confusion metrics
8 print("Confusion Matrix")
9 print(f"Accuracy: {accuracy:.2f}")
10 print(f"Precision: {precision:.2f}")
11 print(f"Recall: {recall:.2f}")
12 print(f"F1-score: {f1_score:.2f}")

```

Confusion Matrix
Accuracy: 0.75
Precision: 0.71
Recall: 0.62
F1-score: 0.67

Gambar 9. Evaluasi *Confusion Matrix* dengan *Python*

Gambar 9 menampilkan hasil evaluasi *confusion matrix* dengan menggunakan *Python* pada *Google Colaboratory* menunjukkan beberapa metrik penting yang digunakan untuk menilai kinerja model. *Accuracy* sebesar 75% menunjukkan bahwa 75% dari seluruh prediksi yang dirancang dengan model adalah benar, baik itu prediksi positif maupun negatif.

Precision sebesar 71.4% menunjukkan bahwa semua prediksi yang dianggap positif oleh model, 71.4% diantaranya benar-benar positif. Dengan kata lain, sekitar 28.6% dari prediksi positif adalah salah. Hal ini dapat menunjukkan seberapa baik model menghindari kesalahan

dalam memprediksi kasus negatif sebagai positif.

Recall sebesar 62.5% menunjukkan bahwa dari semua kasus positif yang sebenarnya terdeteksi oleh model. Dengan kata lain, sekitar 37.5% dari kasus positif yang sebenarnya terlewatkan oleh model.

F1 Score sebesar 67%, memberikan rata-rata harmonis dari *Precision* dan *Recall*. *F1 Score* digunakan untuk menekankan keseimbangan antara kedua metrik. Nilai 67% menunjukkan bahwa model memiliki keseimbangan yang cukup baik antara kemampuan untuk mendeteksi kasus positif dan menghindari kesalahan prediksi kasus negatif sebagai positif.

6. KESIMPULAN

Dalam penelitian ini telah diselesaikan masalah efisiensi proses rekrutmen karyawan baru yang difokuskan pada optimasi penentuan kandidat yang potensial berdasarkan kemiripan profil dan kompetensinya terhadap karyawan. Pendekatan yang dilakukan adalah *collaborative filtering* dengan metode *similarity-based* menggunakan *Jaccard similarity coefficient*. Hasil percobaan menunjukkan keberhasilan memprediksi *ranking* yang dinyatakan dengan skor yang berbeda-beda dan terseleksi berdasarkan *thresholds* tertentu. Implementasi model menggunakan *Google Colaboratory*. Sebagai hasil rekomendasi menggunakan *Jaccard similarity coefficient* didapatkan dari 100 kandidat yang melamar pekerjaan, menjadi 11 kandidat yang memenuhi kebutuhan perusahaan.

DAFTAR PUSTAKA

- [1] A. Fauzi *et al.*, "Pengaruh Kepuasan Kerja dan Stres Kerja Terhadap Turnover Intention Pada Karyawan (Literature Review Manajemen Kinerja)," *Jurnal Ilmu Manajemen Terapan*, vol. 4, no. 2, 2022, doi: 10.31933/jimt.v4i2.
- [2] P. Horodyski, "Applicants' Perception of Artificial Intelligence in The Recruitment Process," *Computers in Human Behavior Reports*, vol. 11, Aug. 2023, doi: 10.1016/j.chbr.2023.100303.
- [3] S. Li, K. Li, and H. Lu, "National Origin Discrimination in Deep-learning-powered Automated Resume Screening," *ArXiv*, Jul. 2023, [Online]. Available: <http://arxiv.org/abs/2307.08624>

- [4] E. Aydın and M. Turan, "An AI-Based Shortlisting Model for Sustainability of Human Resource Management," *Sustainability (Switzerland)*, vol. 15, no. 3, Feb. 2023, doi: 10.3390/su15032737.
- [5] A. S. Pratama, A. Satya Pratama, S. M. Sari, M. F. Hj, M. Badwi, and I. Anshori, "Pengaruh Artificial Intelligence, Big Data Dan Otomatisasi Terhadap Kinerja SDM Di Era Digital," *Jurnal Publikasi Ilmu Manajemen (JUPIMAN)*, vol. 2, no. 4, pp. 108–123, 2023, doi: 10.55606/jupiman.v2i4.2739.
- [6] K. Kant Yadav, H. Kumar Soni, G. Yadav, and M. Sharma, "Collaborative Filtering Based Hybrid Recommendation System Using Neural Network and Matrix Factorization Techniques," *Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE*, vol. 2024, no. 8s, pp. 695–701, 2023, [Online]. Available: www.ijisae.org
- [7] J. Al-Safi and C. Kaleli, "Item genre-based users similarity measure for recommender systems," *Applied Sciences (Switzerland)*, vol. 11, no. 13, Jul. 2021, doi: 10.3390/app11136108.
- [8] S. C. Necula and V. D. Păvăloaia, "AI-Driven Recommendations: A Systematic Review of the State of the Art in E-Commerce," *Applied Sciences (Switzerland)*, vol. 13, no. 9, May 2023, doi: 10.3390/app13095531.
- [9] J. Li and Z. Ye, "Course Recommendations in Online Education Based on Collaborative Filtering Recommendation Algorithm," *Complexity*, vol. 2020, 2020, doi: 10.1155/2020/6619249.
- [10] L. D. Kumalasari and A. Susanto, "Recommendation System of Informatic Technology Jobs using Collaborative Filtering Method Based on LinkedIn Skills Endorsement," *Journal of Information Systems (e-Journal)*, vol. 6, no. 1, 2019, doi: 10.24167/Sisforma.
- [11] I. Tewari and M. Pant, "Artificial Intelligence Reshaping Human Resource Management : A Review," *Proceedings of IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation, ICATMRI 2020*, Dec. 2020, doi: 10.1109/ICATMRI51801.2020.9398420.
- [12] O. Ore and M. Sposato, "Opportunities and Risks of Artificial Intelligence in Recruitment and Selection," *International Journal of Organizational Analysis*, vol. 30, no. 6, pp. 1771–1782, Dec. 2022, doi: 10.1108/IJOA-07-2020-2291.
- [13] F. Fallucchi, M. Coladangelo, R. Giuliano, and E. W. De Luca, "Predicting Employee Attrition using Machine Learning Techniques," *Computers*, vol. 9, no. 4, pp. 1–17, 2020, doi: 10.3390/computers9040086.
- [14] D. Pessach, G. Singer, D. Avrahami, H. Chalutz Ben-Gal, E. Shmueli, and I. Ben-Gal, "Employees Recruitment: A Prescriptive Analytics Approach via Machine Learning and Mathematical Programming," *Decis Support Syst*, vol. 134, Jul. 2020, doi: 10.1016/j.dss.2020.113290.
- [15] Zulkifli, "Analisis Kompetensi Guru Menghadapi Era Revolusi Industri 4.0," *Jurnal Ilmu Sosial dan Pendidikan*, vol. 4, no. 3, 2020, [Online]. Available: <http://ejournal.mandalanursa.org/index.php/IJISIP/index>
- [16] H. Mahmud Nawawi, A. Baitul Hikmah, A. Mustopa, and G. Wijaya, "Model Klasifikasi Machine Learning untuk Prediksi Ketepatan Penempatan Karir," *Jurnal SAINTEKOM*, vol. 14, no. 1, pp. 13–25, Mar. 2024, doi: 10.33020/saintekom.v14i1.512.
- [17] D. Agus Diartono, R. Soelistijadi, H. Februariyanti, and E. Zuliarso, "Model Sistem Rekomendasi Guna Peningkatan Kesesuaian Kebutuhan Program Kampus Merdeka Belajar," *Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, vol. 4, no. 3, pp. 610–621, 2023.
- [18] B. Fernando, A. * Star, and S. Samitha Herath, "Anticipating Human Actions by Correlating Past with The Future with Jaccard Similarity Measures," *Computer Vision and Pattern Recognition*, 2021.
- [19] M. Tang *et al.*, "Evaluating Single-Cell Cluster Stability using The Jaccard Similarity Index," *Bioinformatics*, vol. 37, no. 15, pp. 2212–2214, Aug. 2021, doi: 10.1093/bioinformatics/btaa956.
- [20] A. A. Ternikov, "Skill-Based Clustering Algorithm for Online Job Advertisements," *Izvestiya of Saratov*

- University. Mathematics. Mechanics. Informatics*, vol. 22, no. 2, pp. 250–265, 2022, doi: 10.18500/1816-9791-2022-22-2-250-265.
- [21] M. N. Freire and L. N. de Castro, “e-Recruitment recommender systems: a systematic review,” *Knowl Inf Syst*, vol. 63, no. 1, pp. 1–20, Jan. 2021, doi: 10.1007/s10115-020-01522-8.
- [22] A. Sadikin, H. Yodiansyah, Y. Budiasih, Sugiarti, and I. Henri Kusnadi, “Adaptive Human Resource Management in Confrontation of Globalization’s Challenges,” *Jurnal Ekonomi*, vol. 12, no. 02, p. 2023, 2023, [Online]. Available: <http://ejournal.seaninstitute.or.id/index.php/Ekonomi>
- [23] R. Suwaji, M. Mulyadi, and I. A. Kusuma, “The Role of HR Technology in HR Management Transformation: Improving Organisational Efficiency and Productivity,” *Jurnal Infomasi dan Teknologi*, vol. 6, no. 2, pp. 83–91, 2024, doi: 10.60083/jidt.v6i2.533.
- [24] S. Sudhin and P. Prakash, “An Explorative Study on HRM Based Innovation During the Course of Recruitment,” *J Surv Fish Sci*, vol. 10, no. 2S, pp. 627–638, 2023.
- [25] M. Besta *et al.*, “Communication-Efficient Jaccard similarity for High-Performance Distributed Genome Comparisons,” *Proceedings - 2020 IEEE 34th International Parallel and Distributed Processing Symposium, IPDPS 2020*, pp. 1122–1132, May 2020, doi: 10.1109/IPDPS47924.2020.00118.