



Peningkatan Minat Digital Skill Menggunakan Algoritma K-Medoids Clustering Pada Karyawan

Rizka Zulfiana, Shofa Shofiah Hilabi*, Fitria Nurapriani, Baenil Huda

Fakultas Ilmu Komputer, Sistem Informasi, Universitas Buana Perjuangan Karawang, Karawang

Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat, Indonesia

Email: ¹si20.rizkazulfiana@mhs.ubpkarawang.ac.id, ^{2,*}shofa.hilabi@ubpkarawang.ac.id, ³fitria.apriani@ubpkarawang.ac.id,

⁴baenilhuda88@ubpkarawang.ac.id

Email Penulis Korespondensi: shofa.hilabi@ubpkarawang.ac.id

Submitted: 01/03/2024; Accepted: 30/04/2024; Published: 30/04/2024

Abstrak—Perusahaan Umum Percetakan Uang Republik Indonesia adalah salah satu badan usaha milik negara yang mencetak uang kertas dan dokumen resmi lainnya. Perum Peruri juga berkeinginan untuk mendapatkan wawasan lebih luas mengenai teknologi yang diimplementasikan di dalam perusahaan. Permintaan akan tenaga kerja yang terampil dalam penggunaan teknologi di lingkungan kerja terus meningkat seiring berjalannya waktu. Perum Peruri memiliki 16 kategori Digital Skill, masing-masing kategori tersebut memiliki peminatan yang tinggi, menengah hingga peminatan yang paling rendah. Pada Permasalahan ini data yang diambil adalah belum dilakukan pengelompokan, sehingga kurangnya informasi mengenai jumlah kategori yang memiliki peminatan tertinggi hingga terendah. Dengan menganalisis data peminatan akan membantu menentukan kategori mana yang membutuhkan peningkatan. Kategori dalam peminatan Digital Skill kemudian dapat ditingkatkan dengan menggunakan informasi ini sebagai acuan untuk merancang strategi peningkatan. Penelitian dilakukan dengan menggunakan clustering untuk menentukan jumlah kategori yang diminati oleh karyawan Perum Peruri. Dalam penelitian ini, data penjualan dalam format Excel dianalisis, dan cluster berdasarkan data penjualan produk dibuat menggunakan pendekatan K-Medoids. Informasi penjualan yang diperoleh dari data sekunder yang mengelola peminatan karyawan. Dengan menggunakan RapidMiner, akurasi untuk tiga cluster yang ditetapkan sebagai tertinggi, menengah, dan terendah berdasarkan hasil pengelompokan dipastikan. Cluster pertama dari 16 item yang dianalisis terdiri dari 7 item dengan peringkat tertinggi, cluster kedua memiliki 5 item yang dikategorikan sebagai sedang, dan cluster ketiga memiliki 4 item yang diklasifikasikan sebagai yang terendah. Berdasarkan hasil 4 item peminatan dikategorikan rendah, menunjukkan perlunya pendekatan sosialisasi untuk meningkatkan minat Digital Skill tersebut.

Kata Kunci: Data Mining; Clustering; K-Medoids; Rapidminer; Digital Skill

Abstract—General Company for Printing Money of the Republic of Indonesia is one of the state-owned enterprises that prints banknotes and other official documents. Perum Peruri also wishes to gain more insight into the technology implemented in the company. The demand for a workforce skilled in the use of technology in the work environment continues to increase over time. Perum Peruri has 16 Digital Skill categories, each of these categories has a high, medium to the lowest interest. In this problem, the data taken has not been grouped, so there is a lack of information about the number of categories that have the highest to lowest interest. By analyzing the specialization data, it will help determine which categories need improvement. The categories in Digital Skills specialization can then be improved by using this information as a reference for designing improvement strategies. Research was conducted using clustering to determine the number of categories that Perum Peruri personnel are interested in. In this study, sales data in Excel format was analyzed, and clusters based on product sales data were created using the K-Medoids approach. Sales information obtained from secondary data that manages employee specialization. Using RapidMiner, the accuracy for the three clusters designated as highest, middle, and lowest based on the clustering results was ascertained. The first cluster of 16 items analyzed consisted of 7 items with the highest ranking, the second cluster had 5 items categorized as medium, and the third cluster had 4 items classified as the lowest. Based on the results, 4 items were categorized as low, indicating the need for a socialization approach to increase interest in the Digital Skill.

Keywords: Data Mining; Clustering; K-Medoids; Rapidminer; Digital Skill

1. PENDAHULUAN

Perkembangan dunia teknologi telah mengubah cara berorganisasi, berpikir dan bertindak seiring dengan perkembangan infrastruktur teknologi informasi[1]. Perum Peruri saat ini memiliki wadah riset, penelitian, dan pengembangan sendiri untuk mendukung kesiapan dan kapasitas digital karyawan saat mereka berkembang menjadi mahir. Saat keterampilan digital seseorang dikombinasikan dengan kemampuan kognitif yang kuat, kemampuan memecahkan masalah, dan keahlian bekerja sama, dan berpartisipasi dalam inisiatif digital strategis dalam organisasi[2].

Peminatan dalam keterampilan digital bertujuan untuk memberikan kesempatan kepada karyawan untuk membuat pilihan dan keputusan berdasarkan pemahaman akan potensi diri dan peluang yang ada di bidang keahlian mereka yang luas sangat berguna untuk mencapai tujuan yang diinginkan[3]. Peminatan keterampilan digital sangat penting untuk dimiliki, terutama karena banyak bisnis dan teknologi digital yang mulai beralih ke ranah digital, sehingga karyawan perlu memiliki kemampuan digital yang memadai untuk menciptakan atau mengubah proses bisnis perusahaan[4].

Perum Peruri memiliki komitmen yang tinggi dalam meningkatkan sumber daya manusianya. Khususnya Digital Skill bagi karyawan perusahaan harus dilakukan untuk menyelaraskan dengan kebutuhan perusahaan di bidang yang dipilih, dengan 16 Digital Skill yang tersedia. Hal ini merupakan bagian dari fokus Perum Peruri



dalam mengembangkan bisnis digital dan meningkatkan daya saingnya di industri sekuriti. Data peminatan Digital Skill karyawan Perum Peruri tidak menunjukkan adanya peningkatan pada setiap kategori yang dipilih, sehingga hanya digunakan sebagai arsip dan tidak dimanfaatkan untuk strategi peningkatan peminatan Digital Skill [5].

Untuk meningkatkan kategori, diperlukan pemanfaatan data yang ada, yang dapat menjadi faktor penting dalam menentukan pola peminatan karyawan. Jika data dianalisis dengan efektif, maka informasi mengenai kategori yang paling diminati dan yang kurang diminati dapat terungkap. Dengan demikian, memungkinkan untuk menetapkan strategi peningkatan peminatan pada kategori Digital Skill yang kurang diminati[6]. Untuk mendapatkan hasil yang sesuai, data perlu dikategorikan menjadi tiga kelompok dengan peminatan tertinggi, menengah, dan rendah setelah diolah[7]. Karyawan di Perum Peruri dapat menggunakan data tersebut untuk mendapatkan wawasan tentang kategori yang paling tidak menarik dan kemudian menggunakan temuan tersebut untuk membuat strategi yang akan menarik minat[8].

File Excel yang digunakan dalam penelitian ini berisi rekapitulasi data selama satu bulan. Dengan kemampuan teknologi informasi dalam pemrosesan data yang efisien dan tepat, data ini dapat diolah menjadi sumber informasi berharga[9]. Memproses data dalam jumlah besar dapat dilakukan dengan menggunakan teknik seperti data mining[10]. Kumpulan data yang besar dapat diekstrak untuk mendapatkan informasi, dan data dapat dikelompokkan ke dalam beberapa kelompok berdasarkan karakteristik yang sama dengan menggunakan teknik pengelompokan[11]. Clustering adalah teknik pengelompokan data berdasarkan kesamaan atau perbedaan yang ada di antara data tersebut. Tujuannya adalah untuk membagi dataset ke dalam kelompok-kelompok (cluster), di mana data tersebut berbeda dengan data di seluruh kelompok tetapi memiliki karakteristik yang sama dengan data kelompok lainnya[12]. Ketika memproses kumpulan data berbasis teks dan numerik, teknik K-Medoids yang dikombinasikan dengan metode Euclidean Distance merupakan cara yang efektif untuk menemukan jarak terdekat antara medoids dan titik data lainnya[13]. RapidMiner adalah perangkat lunak pemrosesan data yang dapat melakukan analisis teks, analisis prediktif, dan analisis penambangan data. RapidMiner terkenal sebagai alat yang ampuh untuk mengelola kumpulan data yang berbasis teks dan angka[14].

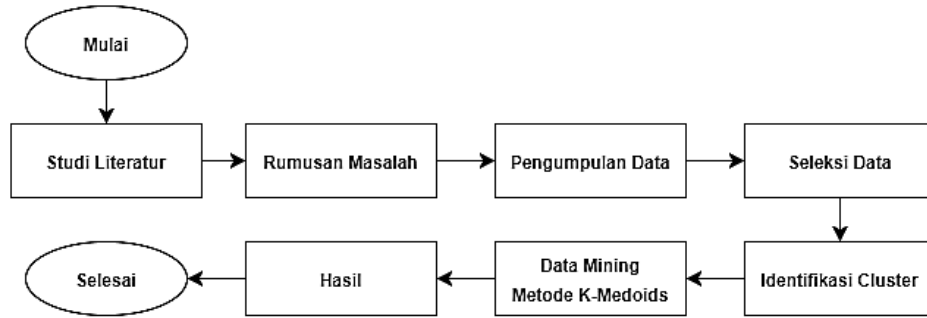
Beberapa penelitian sebelumnya yang telah dilakukan pada tahun 2019 oleh Fitri Hardiyanti, dkk “Penerapan Metode K-Medoids Clustering Pada Penanganan Kasus Diare Di Indonesia” dimana hasil penelitian ini menggunakan data uji dari 34 provinsi, metode K-Medoids digunakan untuk mengatur pengelolaan kasus diare berdasarkan provinsi dengan menggunakan dua cluster. Cluster 1 memiliki nilai rendah sebanyak 31, sedangkan Cluster 2 memiliki nilai tinggi sebanyak 3[15]. Penelitian selanjutnya oleh Desi Asima Silitonga, dkk pada tahun 2019 “Penerapan Metode K-Medoid pada Pengelompokan Rumah Tangga Dalam Perlakuan Memilah Sampah Menurut Provinsi” Hasil yang didapatkan bahwa Proses pemilahan sampah dikelompokkan berdasarkan provinsi dengan menggunakan algoritma K-Medoids. Cluster tingkat pemilahan sampah rendah (C1) dan cluster tingkat pemilahan sampah tinggi (C2) terbentuk dari pengolahan data dalam penelitian ini[16]. Penelitian lainnya oleh Fajar Ageng Bramasta dan Risa Helilintar pada tahun 2021 “Penerapan Data Mining Untuk Menentukan Strategi Penjualan Toko Sepatu” Hasil penelitian berupa tiga kategori yaitu paling populer, rata-rata, dan terendah-digunakan untuk mengelompokkan hasil penelitian cluster[17]. Terdapat juga penelitian oleh Reza Gustrianda dan Dadang Iskandar Mulyana pada tahun 2022 “Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids” Hasil penelitian bahwa kekurangan stok pada produk unggulan dapat diatasi dengan metode pengelompokan. Dengan mengelompokkan produk dengan permintaan tinggi, sedang, dan rendah, perusahaan dapat mengamati hasil dan merencanakan stok dengan lebih efektif[18]. Penelitian terakhir oleh Anggi Octa Fadilah, dkk pada tahun 2023 “Strategi Promosi untuk Meningkatkan Penjualan Kedai Kopi Desimal Menggunakan Algoritma K-Medoids Clustering” Hasil dari penelitian untuk menemukan item menu yang mungkin lebih jarang dijual namun tetap populer, penelitian ini mengklasifikasikan item menu terlaris[19].

Pada peminatan Digital Skill terdapat beberapa kategori seperti Digital Marketing, Internet of Things, Cyber Security dan masih banyak kategori lainnya. Permasalahan dengan data yang diambil adalah belum dilakukan pengelompokan, sehingga kurangnya informasi mengenai jumlah kategori yang memiliki peminatan tertinggi, menengah, hingga terendah. Maka dari itu untuk mengetahui kategori pada peminatan Digital Skill pada peningkatan minat Digital Skill, dilakukan proses clustering pada penelitian ini. Metode clustering yang digunakan adalah algoritma K-Medoids. Dengan melakukan pengelompokan tujuan dari penelitian ini adalah untuk menentukan kategori mana yang memiliki peminatan tertinggi hingga terendah, sehingga dapat membantu dalam peningkatan minat Digital Skill[20]. Algoritma K-Medoids Clustering digunakan dalam penelitian ini, serta bantuan RapidMiner untuk pemrosesan data agar mendapatkan hasil yang akurat ketika K-Medoids Clustering diterapkan. Hasil penelitian ini diharapkan dapat memberikan wawasan tentang sejauh mana minat serta membantu meningkatkan peminatan Digital Skill di kalangan karyawan Perum Peruri.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Dalam metodologi penelitian ini terdiri dari beberapa tahapan yang digunakan untuk mencapai tujuan penelitian yang dilakukan yaitu seperti yang ditunjukkan pada gambar 1 dibawah ini:

**Gambar 1.** Tahapan Penelitian

Gambar 1 merupakan tahapan penelitian yang menjelaskan alur penelitian yang dilakukan untuk mengolah data menggunakan metode K-Medoids yang terdiri dari:

a. Studi Literatur

Pada tahap ini informasi yang terkumpul untuk menghasilkan informasi yang lebih spesifik dengan membandingkan referensi dari berbagai sumber literatur yang relevan diidentifikasi dalam artikel, jurnal, atau sumber online.

b. Rumusan Masalah

Penelitian ini bertujuan untuk mengklasifikasikan minat terhadap keterampilan digital yang diminati dan kurang diminati. Sebagai hasilnya, penelitian ini akan menghasilkan wawasan baru yang mendukung pengolahan data dan perencanaan strategis yang efektif.

c. Pengumpulan Data

Pada tahap ini, Para peneliti menggunakan sumber data sekunder untuk mengumpulkan informasi yang diperlukan untuk penelitian mereka. Dokumen-dokumen berfungsi sebagai sumber data digunakan untuk memberikan data, yang kemudian diproses menjadi file excel yang akan digunakan sebagai sampel dalam penelitian ini[21].

d. Seleksi Data

Hanya data yang sesuai dengan kriteria seleksi yang dapat digunakan. Data seleksi dari Digital 1, Digital 2, dan Digital 3 adalah data yang akan digunakan.

e. Identifikasi Cluster.

Pada tahap ini, dilakukan pemilihan cluster secara acak dari kumpulan data yang telah terkumpul.

f. Pengolahan Data

Data yang terkumpul diproses menggunakan algoritma K-Medoids yang digunakan pada tahap selanjutnya. Perangkat lunak RapidMiner akan digunakan untuk membagi data yang terkumpul menjadi tiga cluster.

g. Hasil

Tahapan ini merupakan langkah terakhir yang akan menghasilkan hasil yang diinginkan dan dapat digunakan untuk mengambil kesimpulan dari penelitian.

2.2 Data Mining

Menemukan pola dan hubungan dalam database yang sangat besar adalah teknik yang dikenal sebagai data mining. Ini adalah proses untuk mengumpulkan, menganalisis, dan mengolah data dari basis data menjadi pengetahuan yang dapat diterapkan oleh para profesional. Untuk mengidentifikasi pola dan hubungan dalam data yang dapat diubah menjadi informasi yang berharga, data mining menggabungkan teknik matematika, statistik, dan informasi [22]. Berbagai industri, termasuk bisnis, perbankan, kesehatan, dan penelitian ilmiah, dapat mengambil manfaat dari data mining. Hal ini dapat membantu bisnis untuk memiliki pemahaman yang lebih dalam tentang data mereka dan memutuskan cara terbaik untuk menggunakannya dengan lebih cerdas[23].

2.3 Algoritma K-Medoids

k-medoids adalah teknik clustering yang menggunakan medoids, bukan rata-rata, untuk merepresentasikan setiap cluster. Medoid adalah objek dalam cluster yang paling mirip dengan rata-rata semua objek dalam cluster. Tujuan dari teknik ini adalah mengembalikan jalur terpendek antara setiap data yang sesuai dengan cluster[24]. Algoritma ini akan mencari lebih banyak cara untuk memperpendek jarak jika jarak terpendek lebih panjang daripada yang lain. Rumus Jarak Euclidean Distance digunakan untuk menentukan jarak antara titik pusat awal dan medoid[25]. Langkah-langkah dasar dari algoritma K-Medoids adalah sebagai berikut:

a. Tentukan jumlah cluster yang diinginkan, yang disebut k.

b. Tentukan k medoid awal secara acak dari data.

c. Masukkan setiap objek ke cluster yang dipimpin oleh medoid terdekat.

d. Hitung medoid baru untuk setiap cluster dengan mencari objek yang paling dekat dengan rata-rata dari semua objek dalam cluster tersebut. Tentukan setiap objek ke cluster terdekat menggunakan rumus perhitungan jarak Euclidean Distance sebagai berikut:



d(x, y) = ||x - y|| = sqrt((sum(xi - yi)^2)/n) = 1 (1)

- e. Masukkan kembali setiap objek ke dalam klaster yang dipimpin oleh medoid terdekat jika medoid yang baru berbeda dengan medoid yang lama.
f. Lakukan langkah 4 dan 5 secara berulang hingga tidak ada perubahan lagi pada medoid.

2.3 Clustering

Clustering adalah teknik menempatkan item bersama-sama sehingga mereka terhubung satu sama lain daripada dengan kelompok penempatan lainnya. Untuk membuat item-item dalam suatu cluster memiliki kemiripan yang tinggi, objek-objek tersebut dikelompokkan ke dalam cluster. Dengan mengelompokkan item atau data, keragaman data dalam suatu kelompok dapat diminimalkan[26].

3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini berasal dari peminatan Digital Skill Perum Peruri yang terdiri dari 16 kategori Digital Skill. Rapidminer akan digunakan untuk mengolah data tersebut dengan menggunakan pendekatan K-Medoids. Oleh karena itu, berikut ini langkah-langkah pengolahan data yang digunakan dalam penelitian ini:

3.1 Seleksi Data

3.1.1 Dataset Awal

Teknik pengelompokan K-Medoids akan digunakan pada tahap pertama pemrosesan data untuk mengolah dataset yang telah dikumpulkan dalam format excel. Langkah ini dilakukan dengan mengumpulkan atau memilih data yang relevan untuk dianalisis berdasarkan pengumpulan data yang sudah ada. Berikut dataset awal ditunjukkan pada tabel 1 di bawah ini.

Tabel 1. Dataset Peminatan Perum Peruri

Table with 5 columns: No, Digital Skill, DS1, DS2, DS3. Rows list 16 digital skills with corresponding values for DS1, DS2, and DS3.

Terdapat 16 Digital Skill yang akan dikelompokkan menggunakan algoritma K-Medoids clustering pada Tabel 1. Kemudian, tabel 2 berisi proses penentuan nilai cluster sebagai titik pusat awal yang akan digunakan.

3.1.2 Menentukan Nilai Clustering

Ketentuan clusterisasi yang diinginkan sebanyak tiga cluster, nilai centroid atau titik tengah dalam penerapan teknik K-Medoids dapat ditentukan secara acak dari data yang tersedia. Pada perhitungan tahap pertama, data dipilih untuk membentuk kelas C1, C2, dan C3 untuk menentukan cluster tingkat tinggi, sedang, dan rendah. Nilai-nilai pusat cluster awal (centroid) ditampilkan pada tabel 2 di bawah ini.

Tabel 2. Nilai Pusat Cluster

Table with 4 columns: Nilai Pusat Cluster, DS1, DS2, DS3. Rows show centroid values for Cloud Computing & Networking, Digital Supply Chain, and Digital Innovation & Collaboration.

Rumus Euclidean Distance akan digunakan untuk menentukan jarak antara setiap pusat cluster dan variabel mengikuti pemilihan acak centroid. Seperti dalam contoh berikut ini:



$$CI(1,1) = \sqrt{(26-18)^2+(11-35)^2+(60-48)^2}=784$$

$$CI(1,2) = \sqrt{(15-18)^2+(31-35)^2+(11-48)^2}=1394$$

$$CI(1,3) = \sqrt{(27-18)^2+(54-35)^2+(37-48)^2}=221$$

Setelah menghitung jarak hasil perhitungan Euclidean Distance ditampilkan dalam tabel 3 berikut:

Tabel 3. Perhitungan Iterasi Pertama

NO	C1	C2	C3	Jarak	Cluster
1	28	59,71599451	82,80700453	28	1
2	37,33630941	48,12483766	94,23905772	37,33630941	1
3	23,72762104	14,86606875	65,55913361	14,86606875	2
4	0	37,17526059	77	0	1
5	45,54119015	36,05551275	32,93933818	32,93933818	3
6	18,22086716	41,97618372	66,88049043	18,22086716	1
7	96,15612305	77,21398837	47,10626285	47,10626285	3
8	77	55,07267925	0	0	3
9	220,4767561	197,8383178	146,904731	146,904731	3
10	65,16133823	48,0416486	36,89173349	36,89173349	3
11	37,17526059	0	55,07267925	0	2
12	33,76388603	63,3245608	105,6645636	33,76388603	1
13	44,24929378	27,27636339	52,43090692	27,27636339	2
14	47,138095	68,2202316	108,4297007	47,138095	1
15	31,41655614	44,14748011	88,55506761	31,41655614	1
16	11,5758369	29,69848481	74,62573283	11,5758369	1
Total Cost		848,7476127		513,4360483	

Tabel 3 menampilkan hasil yang diperoleh dari kedekatan setiap objek selama iterasi pertama. Pada iterasi pertama dihasilkan 8 Digital Skill di dalam Cluster 1, 3 Digital Skill di dalam Cluster 2, dan 5 Digital Skill di dalam Cluster 3. Selanjutnya, lakukan perhitungan ulang pada setiap cluster menggunakan metode yang serupa untuk menentukan nilai centroid. Ikuti langkah yang sama seperti sebelumnya, seperti yang ditunjukkan pada Tabel 4 untuk menghitung jarak setiap objek pada iterasi kedua.

Tabel 4. Nilai Pusat Cluster Baru

Titik Pusat Awal	DS1	DS2	DS3
Data Science & Analytics	32	29	58
Digital Supply Chain	37	65	37
Digital Innovation & Collaboration	87	67	60

Perhitungan centroid terbaru dilakukan menggunakan rumus yang sama yang digunakan pada iterasi pertama. Hasil iterasi kedua adalah dapat dilihat pada tabel 5 dibawah ini.

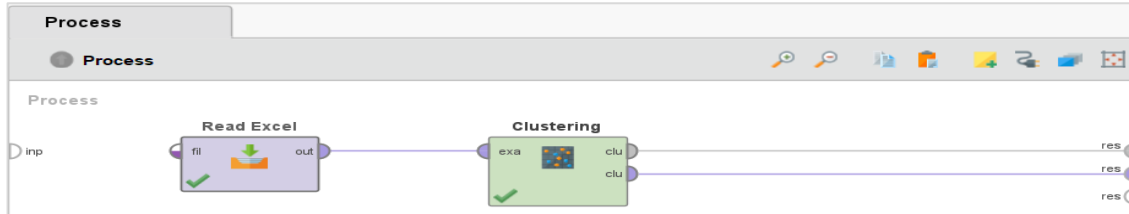
Tabel 5. Perhitungan Iterasi Kedua

NO	C1	C2	C3	Jarak	Cluster
1	19,07878403	59,71599451	82,80700453	19,07878403	1
2	50,019996	48,12483766	94,23905772	48,12483766	2
3	33,03028913	14,86606875	65,55913361	14,86606875	2
4	18,22086716	37,17526059	77	18,22086716	1
5	34,14674216	36,05551275	32,93933818	32,93933818	3
6	0	41,97618372	66,88049043	0	1
7	87,64701934	77,21398837	47,10626285	47,10626285	3
8	66,88049043	55,07267925	0	0	3
9	208,1297672	197,8383178	146,904731	146,904731	3
10	58,05170109	48,0416486	36,89173349	36,89173349	3
11	41,97618372	0	55,07267925	0	2
12	45,3431362	63,3245608	105,6645636	45,3431362	1
13	44,38468204	27,27636339	52,43090692	27,27636339	2
14	56,44466317	68,2202316	108,4297007	56,44466317	1
15	43,02324953	44,14748011	88,55506761	43,02324953	1
16	25,33771892	29,69848481	74,62573283	25,33771892	1
Total Cost		848,7476127		561,5577544	

Tabel 5 menunjukkan hasil perhitungan iterasi kedua. Jika terjadi perubahan dari iterasi pertama hingga iterasi kedua, lakukan perhitungan ulang hingga diperoleh hasil sesuai pada iterasi tersebut. Selanjutnya, data diolah menggunakan RapidMiner untuk menganalisis dan pengujian.

3.2 Proses Clustering Rapidminer

RapidMiner Studio digunakan untuk menganalisis dan melakukan pengujian dengan tujuan membentuk kelompok cluster. Ada beberapa proses pengolahan dataset yang dapat dilakukan untuk menentukan nilai cluster dengan tingkat data tertinggi, menengah, dan rendah menggunakan metode K-Medoids yang ditunjukkan pada gambar 2 dibawah ini.



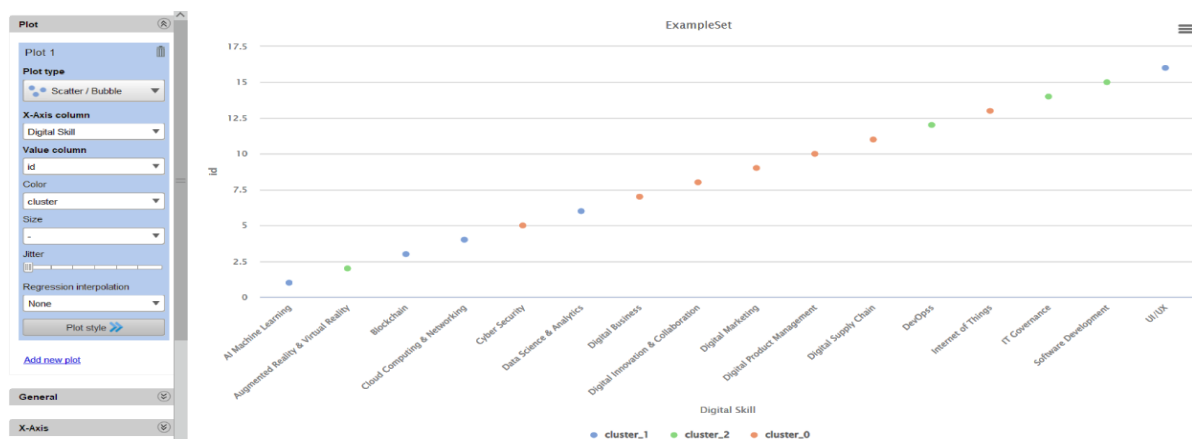
Gambar 2. Proses Clustering Rapidminer

Tahapan awal di jelaskan pada gambar 2 proses awal memasukan data excel kedalam Rapidminer menggunakan operator Read Excel. Selanjutnya gunakan algoritma K-Medoids untuk menentukan jumlah cluster yang diinginkan, lalu masukan 3 nilai yang akan dikelompokan menjadi 3 cluster. Gambar 3 menampilkan hasil dari proses clustering.

Row No.	id	Digital Skill	cluster	DS1	DS2	DS3
1	1	AI Machine Learning	cluster_1	26	11	60
2	2	Augmented Reality & Virtual Reality	cluster_2	15	31	11
3	3	Blockchain	cluster_1	27	54	37
4	4	Cloud Computing & Networking	cluster_1	18	35	48
5	5	Cyber Security	cluster_0	61	47	57
6	6	Data Science & Analytics	cluster_1	32	29	58
7	7	Digital Business	cluster_0	76	94	97
8	8	Digital Innovation & Collaboration	cluster_0	87	67	60
9	9	Digital Marketing	cluster_0	233	83	57
10	10	Digital Product Management	cluster_0	57	77	79
11	11	Digital Supply Chain	cluster_0	37	65	37
12	12	DevOpss	cluster_2	2	13	28
13	13	Internet of Things	cluster_0	35	73	63
14	14	IT Governance	cluster_2	11	8	10
15	15	Software Development	cluster_2	19	30	17
16	16	UI/UX	cluster_1	21	40	38

Gambar 3. Hasil proses Clustering pada Rapidminer

Gambar 3 di atas menampilkan data yang berhasil diproses. Visualisasi data anggota kelompok dapat terlihat setelah proses clustering, pilih parameter model Scatter/Bubble untuk mendapatkan hasil visualisasi yang terlihat seperti gambar 4 dibawah ini.



Gambar 4. Visualisasi Cluster Scatter/Bubble

Pada Rapidminer visualisasi model cluster menggunakan Scatter/Bubble direpresentasikan terdapat 16 kategori yang diberi nama Cluster_0, Cluster_1, dan Cluster_2, seperti yang terlihat pada gambar 4. Kelompok cluster tertinggi yaitu Internet of Things terdapat pada Cluster_0. Kelompok cluster menengah yaitu UI/UX

terdapat pada Cluster_1. Kelompok cluster terendah yaitu Software Development terdapat pada Cluster_2. Meskipun cara menetapkan nilai awal cluster berbeda dengan proses perhitungan manual menggunakan alat Rapidminer, cluster dan kategori kelompok ini menghasilkan kesamaan dengan perhitungan secara manual. Hasil antara Rapidminer dan hasil cluster manual tidak jauh berbeda setelah memasukkan 16 dataset Digital Skill, gambar 5 berikut ini menampilkan hasil nilai model cluster:

Cluster Model

```
Cluster 0: 7 items
Cluster 1: 5 items
Cluster 2: 4 items
Total number of items: 16
```

Gambar 5. Hasil Cluster RapidMiner

Gambar 5 menampilkan bahwa perhitungan hasil cluster dari 16 dataset menghasilkan tiga cluster, dihasilkan Cluster 0 dengan jumlah item 7, Cluster 1 dengan jumlah item 5, dan Cluster 2 dengan jumlah 4 item. Hasil akhir menunjukkan bahwa pengelompokan sudah tepat setelah dilakukan perhitungan manual dan menggunakan aplikasi Rapidminer.

Tabel 6. Hasil Kedekatan Perhitungan Manual

Cluster	Anggota Cluster	Jumlah Anggota
C1	CS, DB, DIC, DM, DPM, DSC, IOT	7
C2	AIM, BOC, CCN, DSA, UIUX	5
C3	ARVR, DOP, ITG, SWD	4

Tabel 6 di atas menampilkan hasil clustering dengan jumlah anggota cluster dan jumlah anggota. Berikut perhitungan menggunakan Algoritma K-Medoids untuk 16 item data dengan alat bantu RapidMiner. Dalam Tabel 6, terdapat 4 item Digital Skill yang memiliki minat paling rendah, yaitu Augmented Reality & Virtual Reality, DevOps, IT Governance, dan Software Development. Dari hasil tersebut, beberapa kategori Digital Skill memerlukan strategi undangan atau ajakan untuk meningkatkan minat, sehingga kategori-kategori tersebut tetap bersaing secara kompetitif.

4. KESIMPULAN

Hasil pengelompokan kategori menggunakan algoritma K-Medoids dengan menggunakan alat Rapidminer menunjukkan bahwa, peminatan Digital Skill memerlukan strategi undangan atau ajakan kepada karyawan Perum Peruri khususnya yang memiliki kategori Digital Skill yang rendah. Penerapan algoritma K-Medoids dapat mengklasifikasikan kategori yang paling banyak diminati, termasuk mampu mengidentifikasi kategori yang kurang diminati. Hasil ini dapat membantu peminatan Digital Skill dalam meningkatkan strategi ajakan dan meningkatkan minat karyawan. 7 item kategori berada di cluster 0, 5 item kategori berada di cluster 1, dan 4 item kategori berada di cluster 2, sesuai dengan hasil pengelompokan yang diperoleh dengan menggunakan algoritma K-medoids. Dengan menampilkan 3 clustering dalam pengujian Rapidminer ini lebih menunjukkan bahwa tingkat ketepatan clustering yang sama.

REFERENCES

- [1] S. S. Hilabi, "TechnoXplore Jurnal Ilmu Komputer & Teknologi Informasi ISSN: 2503-054X Vol 4 No: 1, April 2019," *J. Ilmu Komput. Teknol. Inf.*, vol. 4, no. 1, pp. 28–37, 2019.
- [2] L. A. Adha, "Digitalisasi Industri Dan Pengaruhnya Terhadap Ketenagakerjaan Dan Hubungan Kerja Di Indonesia," *J. Kompil. Huk.*, vol. 5, no. 2, pp. 267–298, 2020, doi: 10.29303/jkh.v5i2.49.
- [3] D. Gustian and M. S. Al-Farits, "Data Mining Untuk Melihat Minat Belajar Siswa Menerapkan Metode K-Means," *J. Inf. Syst. Res.*, vol. 4, no. 3, pp. 775–784, 2023, doi: 10.47065/josh.v4i3.3218.
- [4] A. I. Rasidi and N. G. Ramadhan, "Perancangan Website Company Profile Perusahaan CV. CUP10INDO Menggunakan Metode Design Thinking," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 5, pp. 457–465, 2023, [Online]. Available: <https://djournals.com/klik>
- [5] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.
- [6] S. N. Adila and N. Aziz, "Pengaruh Strategi Promosi terhadap Keputusan Pembelian yang dimediasi oleh minat beli pada konsumen Restoran KFC cabang khatib Sulaiman Padang," 2019.
- [7] A. Z. Kamalia, E. A. Pradana, and N. Surojudin, "Penerapan Algoritma K-Means Dalam Klasterisasi Penjualan Laptop," *J. SIGMA*, vol. 1, no. 1, pp. 147–158, 2022, [Online]. Available: <https://jurnal.pelitabangsa.ac.id/index.php/sigma/article/view/1440>
- [8] S. F. Mulaki, N. Setiyawati, and A. F. Wijaya, "Analisis Data Mahasiswa Menggunakan Algoritma K-Means Clustering



- sebagai Dasar Pelaksana Promosi,” *JBASE - J. Bus. Audit Inf. Syst.*, vol. 1, no. 2, pp. 30–39, 2018, doi: 10.30813/v1i2.1259.
- [9] A. Agneresa, A. L. Hananto, S. S. Hilabi, A. Hananto, and T. Tukino, “Strategi Promosi Penerapan Data Mining Mahasiswa Baru Dengan Metode K-Means Clustering,” *Dirgamaya J. Manaj. dan Sist. Inf.*, vol. 2, no. 2, pp. 25–34, 2022, doi: 10.35969/dirgamaya.v2i2.275.
- [10] R. Adrianto and A. Fahmi, “Penerapan Metode Clustering Dengan Algoritma K-Means Untuk Rekomendasi Pemilihan Jalur Peminatan Sesuai Kemampuan Pada Progam Studi Teknik Informatika - S1 Universitas Dian Nuswantoro,” *JOINS (Journal Inf. Syst.*, vol. 1, no. 2, pp. 101–116, 2019, [Online]. Available: <http://publikasi.dinus.ac.id/index.php/joins/article/view/1302>
- [11] A. Lia Hananto et al., “Analysis of Drug Data Mining with Clustering Technique Using K-Means Algorithm,” *J. Phys. Conf. Ser.*, vol. 1908, no. 1, 2021, doi: 10.1088/1742-6596/1908/1/012024.
- [12] A. Aditya, I. Jovian, and B. N. Sari, “Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019,” *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 51, 2020, doi: 10.30865/mib.v4i1.1784.
- [13] S. Nuraini, I. Gunawan, and W. Saputra, “Utilization of K-Medoids Algorithm for Klustering of Oil Palm Sprouts,” *JOMLAI J. Mach. Learn. Artif. Intell.*, vol. 1, no. 1, pp. 11–22, 2022, doi: 10.55123/jomlai.v1i1.160.
- [14] A. E. Wijaya, R. Bani, S. Sukarni, and S. A. Weighting, “Jurnal Teknologi Informasi dan Komunikasi STMIK Subang, Oktober 2019 ISSN: 2252-4517,” no. April, pp. 100–110, 2019.
- [15] F. Hardiyanti, H. S. Tambunan, and I. S. Saragih, “Penerapan Metode K-Medoids Clustering Pada Penanganan Kasus Diare Di Indonesia,” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 598–603, 2019, doi: 10.30865/komik.v3i1.1666.
- [16] D. A. Silitonga, A. P. Windarto, D. Hartama, and Sumarno, “Penerapan Metode K-Medoid pada Pengelompokan Rumah Tangga Dalam Perlakuan Memilah Sampah Menurut Provinsi,” *Semin. Nas. Sains Teknol. Inf. SENSASI 2019 ISBN*, vol. 2, pp. 313–318, 2019.
- [17] S. Nasional, I. Teknologi, F. A. Bramasta, R. Helilantar, T. Informatika, and F. Teknik, “Penerapan Data Mining Untuk Menentukan Strategi Penjualan Toko Sepatu,” pp. 236–241, 2021.
- [18] R. Gustrienda and D. I. Mulyana, “Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids,” *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 27, 2022, doi: 10.30865/mib.v6i1.3294.
- [19] A. Octa Fadhilah, B. Huda, and A. Hananto, “Strategi Promosi untuk Meningkatkan Penjualan Kedai Kopi Desimal Menggunakan Algoritma K-Medoids Clustering,” *J. Ris. Komputer*, vol. 10, no. 1, pp. 2407–389, 2023, doi: 10.30865/jurikom.v10i1.5561.
- [20] Didik Maulana and S. Sundari, “Penerapan Algoritma K-Medoids Dalam Klasterisasi Penyebaran,” no. 1, pp. 51–59, 2022.
- [21] T. Widyanti, S. S. Hilabi, A. Hananto, Tukino, and E. Novalia, “Implementasi K-Means dan K-Nearest Neighbors pada Kategori Siswa Berprestasi,” *J. Inf. dan Teknol.*, vol. 5, no. 1, pp. 75–82, 2023, doi: 10.37034/jidt.v5i1.255.
- [22] D. Jollyta, W. Ramdhan, and M. Zarlis, *Konsep Data Mining Dan Penerapan*. Deepublish, 2020.
- [23] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*. Deepublish, 2020.
- [24] R. K. Purba and E. Bu’ulolo, “Implementasi Algoritma K-Medoids dalam Pengelompokan Mahasiswa yang Layak Mendapat Bantuan Uang Kuliah Tunggal,” *INSOLOGI J. Sains dan Teknol.*, vol. 1, no. 2, pp. 79–86, 2022, doi: 10.55123/insologi.v1i2.195.
- [25] P. A. Kusuma and A. U. Firmansyah, “Deteksi Penyebaran Penyakit Tuberkulosis dengan Algoritma K-Means Clustering Menggunakan Rapid Miner,” *J. Teknol. Inform. dan Komput.*, vol. 8, no. 2, pp. 41–54, 2022, doi: 10.37012/jtik.v8i2.1173.
- [26] C. Nas, “Data Mining Pengelompokan Bidang Keahlian Mahasiswa Menggunakan Algoritma K-Means (Studi Kasus: Universitas Cic Cirebon),” *Syntax J. Inform.*, vol. 9, no. 1, pp. 1–14, 2020.