

Klasterisasi Data Pencari Kerja di Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul Menggunakan Algoritma K-Means

Desy Susilowati^a, Yanuar Wicaksono^{b*}

^{a,b} Program Studi Sistem Informasi/Universitas Alma Ata, Yogyakarta 55184, Indonesia

Informasi Naskah:

Diterima: 31 Januari 2024/ Direview: 31 Januari 2024/ Direvisi: 21 Februari 2024/ Disetujui Terbit: 01 Maret 2024

DOI: 10.33369/pseudocode.11.2.54-58

*Korespondensi: yanuar@almaata.ac.id

Abstract

Job seekers applying for employment must adhere to the policies of the desired company. A current and relevant issue is the high number of job seekers who struggle to secure employment due to various factors, including educational levels and job availability in different regions. Job applicants across various companies often possess diverse age and educational qualifications, particularly in the Bantul Regency, D.I Yogyakarta. The data utilized in this research is sourced from the Bantul District Manpower and Transmigration Office. Data processing is carried out using Data Mining or Knowledge Discovery in Database, employing the clustering method with the K-Means algorithm to categorize job seeker data into desired clusters. The aim of this study is to group job seeker data in the Bantul Regency and segregate them into groups with distinct characteristics. Testing is conducted using the RapidMiner software, a Data Mining tool with various available methods. The grouping of job seeker data is expected to assist the government in evenly distributing jobs to job seekers and reducing the accumulation of job seeker data annually. Additionally, it can aid the Bantul District Manpower and Transmigration Office in planning and allocating resources according to needs. The implementation results of the K-Means algorithm on job seeker data, tested from $k=2$ to $k=10$, yield the best cluster at $k=2$ with a Davies Bouldin Index (DBI) value of 0.523. Cluster_0 consists of 214 members, while Cluster_1 consists of 36 members. Therefore, the smaller the Davies Bouldin Index (DBI) value, the higher the similarity level of data within one cluster.

Keywords: Data Mining, Job Seekers, RapidMiner, Clustering, K-Means Algorithm

1. Pendahuluan

Pencari kerja merupakan angkatan kerja yang belum mendapatkan pekerjaan tetap dan menganggur [1]. Pencari kerja saat ini sering mengalami kesulitan dalam melamar pekerjaan akibat beberapa faktor, termasuk tidak memenuhi persyaratan tingkat pendidikan dan ketersediaan lapangan pekerjaan yang belum merata di berbagai daerah [2]. Pencari kerja mempunyai hak untuk mendapatkan kesempatan bekerja sesuai dengan keahlian dan keterampilan. Pencarian pekerjaan adalah tantangan yang tidak hanya dihadapi oleh individu pencari kerja, tetapi juga oleh perusahaan yang mencari calon pekerja yang sesuai [3]. Pencari kerja saat ini dapat mengakses berbagai platform *online* atau mengunjungi kantor Dinas Tenaga Kerja dan Transmigrasi di wilayah masing-masing yang menyediakan informasi tentang lowongan pekerjaan, sehingga memudahkan mereka untuk mencari pekerjaan yang sesuai dengan kualifikasi dan preferensi mereka.

Menurut data pada laman resmi Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul [4] tahun 2020 sebesar 4816, tahun 2021 sebesar 6175, tahun 2022 sebesar 6762, dan tahun 2023 sebesar 3787. Jumlah pencari kerja mengalami peningkatan pada jumlah yang mencari lowongan pekerjaan di instansi tersebut sebagaimana ditunjukkan pada gambar 1. Hal ini menyebabkan penumpukan data yang terus menerus, dan menghambat akses cepat terhadap informasi [2]. Data pencari kerja juga dapat dianalisis untuk menghasilkan pengetahuan

yang berguna bagi Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul dalam memahami pola dan karakteristik dari pencari kerja, sehingga dapat membantu dalam perencanaan dan alokasi sumber daya untuk pelatihan atau pengembangan keterampilan sesuai dengan kebutuhan.



Gambar 1. Jumlah Pencari Kerja

Analisis klasterisasi membagi data berdasarkan kesamaan karakteristik kelompok-kelompok pencari kerja untuk mempermudah dalam mengekstrak pengetahuan data pencari kerja. Informasi ini dapat berguna dalam perencanaan sumber daya manusia secara lebih akurat.

Analisis data menggunakan teknologi *data mining* yang telah digunakan secara luas untuk menggali data dari *database* berukuran besar yang memiliki tingkat kerumitan yang tinggi [5]. Terdapat tiga aspek utama dalam *data mining*, yaitu *classification*, *association*, dan *clustering*. Di antara ketiga

aspek tersebut, *clustering* fokus pada pengelompokan data ke dalam kategori objek yang memiliki kesamaan berdasarkan subjek penelitian. *Clustering* adalah proses mengeksplorasi atau penjelasan data tanpa adanya model atau hipotesis yang telah ditetapkan sebelumnya, tetapi bertujuan untuk memahami ciri umum atau struktur data yang memiliki dimensi tinggi melalui pengelompokan [6]. Algoritma *clustering K-Means* adalah salah satu metode yang sederhana untuk dijalankan dan memiliki kapabilitas untuk mengelompokkan data dalam jumlah besar dengan waktu komputasi yang cepat dan efisien [7].

K-Means adalah algoritma *clustering* yang menggunakan jarak sebagai dasar untuk membagi data ke dalam beberapa *cluster*, dan hanya berlaku untuk atribut numerik. Algoritma ini mengelompokkan data ke dalam *cluster* sehingga data dengan karakteristik serupa ditempatkan dalam satu kelompok, sementara data yang memiliki karakteristik yang berbeda ditempatkan dalam kelompok yang berbeda [6]. Algoritma *K-Means* ini bersifat fleksibel dalam menentukan jumlah *cluster* yang akan dibuat [8]. Dalam menggunakan algoritma *K-Means*, pengguna harus memiliki jumlah *cluster* yang tepat, yaitu sebanyak k . Banyaknya kelompok yang dapat dikenali bergantung pada jumlah serta variasi objek data. Dengan bersifat fleksibel mengakibatkan penentuan jumlah *cluster* menjadi bias karena sudut pandang yang berbeda-beda terhadap pemahaman data. Beberapa metode untuk menentukan nilai k , yaitu jumlah *cluster* yang dibentuk secara dinamis, salah satunya adalah metode *Davies Bouldin Index (DBI)* [2].

Penerapan algoritma *K-Means* dalam penelitian ini didasarkan pada beberapa pertimbangan. Pertama, *K-Means* merupakan algoritma *clustering* yang umum digunakan dan terbukti efektif dalam mengelompokkan data. Kedua, algoritma ini relatif cepat dan efisien, sehingga cocok untuk menangani jumlah data yang cukup besar. Ketiga, *K-Means* memiliki sifat yang mudah diinterpretasi, di mana hasilnya dapat dimengerti dengan jelas oleh pihak yang tidak memiliki latar belakang teknis yang mendalam. Keempat, algoritma ini memiliki fleksibilitas dalam menangani berbagai jenis data, termasuk data pencari kerja yang memiliki karakteristik yang beragam [9].

DBI adalah suatu metode evaluasi internal yang memvalidasi sejauh mana pengelompokan dilakukan berdasarkan kuantitas dan fitur yang terkandung dalam *dataset* [10]. DBI menggambarkan tingkat kesamaan antar *cluster*, dengan asumsi bahwa *cluster* tersebut memiliki kepadatan data yang mengalami penurunan jarak dari vektor karakteristiknya. Dari perhitungan *Davies Bouldin Index (DBI)* dapat disimpulkan bahwa jika semakin kecil nilai *Davies Bouldin Index (DBI)* yang diperoleh (non negatif ≥ 0) semakin baik kualitas *cluster* tersebut.

DBI digunakan untuk menentukan jumlah klaster optimal dalam proses pengelompokkan. Fungsinya adalah untuk mengukur dan memaksimalkan jarak antar-klaster, sehingga klaster satu dengan yang lain memiliki perbedaan yang maksimal. Secara esensial, evaluasi DBI berfokus pada perbandingan antara klaster ke- k dan ke- k' , di mana nilai harus maksimal untuk perbedaan dalam klaster dan minimal untuk kesamaan di dalam klaster. DBI menunjukkan korelasi positif untuk kasus di dalam klaster dan korelasi negatif untuk kasus antar klaster [11].

Penerapan *K-Means* dalam menganalisis data pencari kerja diharapkan agar memberikan dampak positif terhadap operasional dan kebijakan Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul, dengan pembagian data berdasarkan jumlah *cluster* yang telah optimal. Jumlah kelompok yang dapat dibentuk dari para pencari kerja yang belum mendapatkan pekerjaan berdasarkan umur, tingkat pendidikan, jenis kelamin, status perkawinan, dan keterampilan.

2. Metodologi Penelitian

Penelitian ini menggunakan metode *Knowledge Discovery in Database (KDD)* untuk pendekatan, cakupan atau objek, definisi operasional variabel/deskripsi fokus penelitian, lokasi, populasi dan sampel/informan, bahan dan alat utama, teknik pengumpulan data, dan teknik analisis data. Proses KDD, yang melibatkan sejumlah tahapan yang meliputi pemilihan data, pembersihan data, integrasi data, transformasi data, evaluasi pola, hingga presentasi pengetahuan [12].

2.1. Selection Data

Tahap pertama adalah pengambilan data yang akan digunakan untuk penelitian ini. Data *selection* memilih data yang siap digunakan dalam analisis data, terdapat *value* data yang tidak lengkap atau kosong sehingga tidak dapat digunakan dalam analisis. Dari 10.549 yang dapat digunakan hanya 250 data dengan kelengkapan data yang diharapkan.

Selection data merupakan tahap dilakukannya analisis terhadap data yang relevan dalam *database*. Hal ini diperlukan karena tidak semua data yang dibutuhkan dalam proses data *mining*. Dalam data pencari kerja, hanya terdapat lima atribut yang digunakan, yaitu umur dengan tipe data *numerik*, tingkat pendidikan dengan tipe data *polynomial*, jenis kelamin dengan tipe data *polynomial*, status perkawinan dengan tipe data *polynomial* dan keterampilan dengan tipe data *polynomial*.

Tipe data *polynomial* atau dalam analisis data, variabel *polynomial* sering dianggap sebagai salah satu dari tipe data kategorikal adalah salah satu tipe data yang digunakan untuk merepresentasikan variabel yang memiliki kategori atau label yang terstruktur, tetapi tidak memiliki urutan atau tingkat peringkat tertentu di antara kategori-kategori tersebut. Variabel *polynomial* dapat mengambil nilai-nilai dari sekumpulan kategori yang diskrit dan tidak terurut.

2.2. Preprocessing Data

Tahap kedua adalah tahap *preprocessing* data, yaitu menstandarkan nilai data yang diinputkan oleh pencari kerja. Di bawah ini merupakan tabel pengkoreksian nilai data.

Tabel 1. Koreksi Data

Nilai Data	Koreksi Data
SMAN	SMA/MA
KAWIN	NIKAH
MENJAHIT	ADA
KOMPUTER	ADA

2.3. Transformation Data

Langkah ketiga dalam pengolahan *dataset* melibatkan transformasi awal untuk memungkinkan RapidMiner menganalisis data yang akan diproses. Oleh karena itu, perlu dilakukan normalisasi pada atribut tingkat pendidikan, status perkawinan, jenis kelamin, dan keterampilan yang awalnya berformat *string*, menjadi format numerik agar dapat menjalani proses data *mining* menggunakan aplikasi RapidMiner, seperti yang ditunjukkan dalam Tabel 2 Normalisasi pada atribut tingkat pendidikan sampai dengan Tabel 5 Normalisasi pada atribut keterampilan di bawah ini.

Tabel 2. Normalisasi pada atribut tingkat pendidikan

Tingkat Pendidikan	Inisialisasi
DIPLOMA I/II/III/IV	1
PROFESI	2
TIDAK/BELUM TAMAT SD	3
SD	4
SMP	5
SMA/MA	6
SMK	7
STRATA I/II/III	8

Tabel 3. Normalisasi pada atribut jenis kelamin

Jenis Kelamin	Inisialisasi
LAKI-LAKI	1
PEREMPUAN	2

Tabel 4. Normalisasi pada atribut status perkawinan

Status Perkawinan	Inisialisasi
BELUM NIKAH	1
NIKAH	2
CERAI HIDUP	3
CERAI MATI	4

Tabel 5. Normalisasi pada atribut keterampilan

Keterampilan	Inisialisasi
TIDAK ADA	1
ADA	2

Setelah dilakukan proses transformation data dengan menggunakan kode normalisasi pada tabel di atas, kemudian terjadi perubahan *type* pada atribut tingkat pendidikan, jenis kelamin, status perkawinan dan keterampilan yang awalnya *polynomial*, menjadi numerik karena proses normalisasi.

2.4. Data Mining

Langkah keempat melakukan pengelompokan data pencari kerja menggunakan teknik data mining yang disebut Algoritma K-Means. Langkah-langkah dalam proses pengelompokan ini menggunakan metode K-Means adalah sebagai berikut [13]:

- Menentukan jumlah *cluster* yang diinginkan.

Jumlah *cluster* ini menentukan seberapa banyak kelompok data yang akan dihasilkan dalam proses *clustering*.

- Menetapkan titik pusat awal (*centroid*).

Centroid awal diberikan secara acak dan jumlahnya sesuai dengan jumlah *cluster* yang ditentukan sebelumnya. *Centroid* awal ini menjadi titik pusat pertama atau awal dari setiap *cluster*.

- Menghitung jarak dari setiap objek data ke titik pusat (*centroid*).

Setelah titik pusat telah ditetapkan, langkah berikutnya adalah mengukur jarak setiap data dengan titik pusat yang telah ditetapkan sebelumnya menggunakan metode *Euclidean Distance*. Rumus yang digunakan untuk menghitung jarak dari setiap objek data ke titik pusat adalah sebagai berikut:

$$\text{Euclidean Distance (d)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan:

x : nilai data

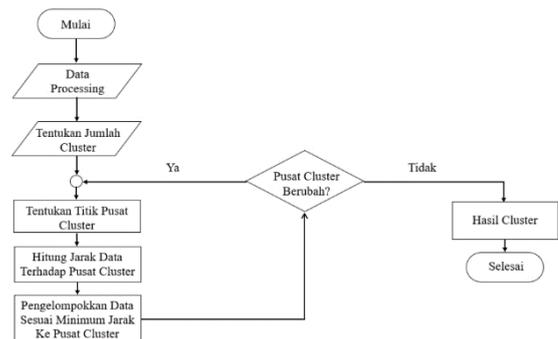
y : nilai *centroid*

i : variabel ke - i

n : jumlah variabel

- Mengelompokkan data ke dalam *cluster* dengan jarak terdekat dari *centroid* sebagai *cluster* yang sesuai.

- Melakukan pengecekan, jika posisi *cluster* berubah maka mengulangi langkah ketiga dengan nilai *centroid* baru, dan jika *cluster* tidak berubah atau memiliki persamaan dengan *cluster* sebelumnya maka proses perhitungan dihentikan dan *cluster* tersebut dicatat sebagai hasil *cluster* yang terbentuk.



Gambar 1 Algoritma K-Means

Dalam proses pengolahan data, dilakukan pembentukan model untuk *clustering*, dengan menggunakan aplikasi RapidMiner [14].

2.5. Interpretation/Pembahasan

Interpretasi dalam data *mining* melibatkan proses analisis dan penafsiran hasil yang diperoleh dari algoritma *k-means*. Ini mencakup langkah-langkah evaluasi performa model dan interpretasi temuan yang dihasilkan dari proses analisis data. Dalam tahap ini, dilakukan pengukuran sejauh mana model data *mining* dapat memprediksi atau mengklasifikasikan data dengan akurasi tinggi.

3. Hasil dan Pembahasan

3.1. Hasil

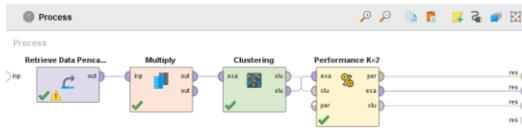
Setelah dilakukan proses transformasi data pada data pencari kerja, hasilnya dapat dilihat pada Gambar 4 di bawah ini.

no	nama	jenis kelamin	status perkawinan	tingkat pendidikan	keterampilan	umur
1	AGUS WIDODO	2	4	5	2	21
2	ERI KURNIAWAN	2	3	7	1	30
3	BINTA AULIA	1	3	3	2	23
4	ADIK TRI UNTORO	2	3	4	2	31
5	ARYATAMA BAKHIS RAMADAHAN	2	4	5	2	21
6	EDI TAMA	2	4	5	2	21
7	ANIAS SEPTIAN MUKTI KURNIA	2	4	5	2	23
8	ADHA FEBRIANTO	2	4	5	2	21
9	IMAWAN JATI PRABOWO	2	4	7	2	27
10	RISKA FARIDA	1	4	4	2	26
11	GALIH ROSO TUNGGAL JATI	2	4	5	2	21
12	MAYA IKA RAHMAWATI	1	3	6	2	28

Gambar 2 Data Pencari Kerja

Pada pengelompokan data pencari kerja, proses data mining dilakukan menggunakan perangkat lunak RapidMiner. Langkah pertama sebelum melaksanakan pengelompokan adalah mengimpor data, di mana data yang diimpor merupakan data *excel* yang telah dinormalisasi.

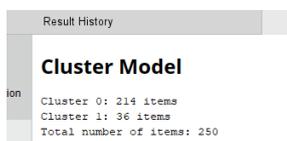
Setelah berhasil diimpor ke RapidMiner, letakkan *dataset* tersebut dari *repository* ke bagian proses dengan *click* dan *drag*. Gambar 6 merupakan susunan operator di aplikasi RapidMiner dalam proses *clustering* menggunakan Algoritma *K-Means*.



Gambar 3 Proses Clustering

Sebelum memulai proses Data Mining, penjelasan berikut ini memberikan gambaran tentang serangkaian operator yang digunakan dalam perangkat RapidMiner untuk melakukan *clustering* data pencari kerja menggunakan Algoritma *K-Means*.

- Operator *Retrieve* : untuk mendapatkan akses ke data pencari kerja yang disimpan dalam penyimpanan lokal dan memuatnya ke dalam proses.
- Operator *Multiply* : untuk mempermudah pengujian ganda agar tidak perlu memasukkan operator *Read csv* kembali.
- Operator *Clustering* : untuk mengelompokkan data ke dalam kelompok atau cluster yang memiliki karakteristik atau ciri-ciri yang sama.
- Operator *Performance* : untuk memilih model atau algoritma yang paling sesuai untuk digunakan dalam proses analisis data [15].



Gambar 4 Hasil Clustering

Dan setelah melalui proses pengelompokan, hasilnya

terlihat seperti yang ditampilkan dalam Gambar 7. menunjukkan pembentukan 2 *cluster* (K2), dengan *cluster_0* memiliki 214 anggota dan *cluster_1* memiliki 36 anggota.

3.2. Pembahasan

Langkah *clustering* dengan algoritma *K-Means* dilakukan berulang dengan parameter $k=2$ hingga $k=10$. Mengevaluasi hasil pengelompokan dengan menggunakan *Davies Bouldin Index* (DBI). Hasil dari penerapan Algoritma *K-Means* pada data pencari kerja untuk menentukan nilai DBI dalam proses data mining dengan RapidMiner dapat dilihat pada gambar 8.



Gambar 5 Nilai Davies Bouldin Index

Hasil akhir dari pengelompokan *cluster* terbaik DBI dapat ditemukan pada Tabel 6. Penerapan Algoritma *K-Means* pada data pencari kerja menghasilkan bahwa nilai k optimum terletak pada K2 dengan nilai DBI sebesar 0.523.

Tabel 6. Hasil DBI

Cluster	DBI
K2	0.523
K3	0.635
K4	0.657
K5	0.636
K6	0.692
K7	0.781
K8	0.708
K9	0.735
K10	0.833

Dengan mendapatkan nilai optimum k terletak pada K2 dengan *centroid* yang dihasilkan di dalam *cluster* 0 yaitu jenis kelamin Laki-laki, status perkawinan Belum Kawin atau Belum Menikah, tingkat pendidikan SMK, keterampilan Ada dan umur 23, sedangkan pada *cluster* 1 jenis kelamin Laki-laki, status perkawinan Kawin atau Sudah Menikah, tingkat pendidikan SMK, keterampilan Ada dan umur 38.

Centroid terakhir dari *clustering K-Means* pencari kerja dapat dilihat pada tabel 7 di bawah ini.

Tabel 7. Centroid Terakhir

Atribut	Cluster_0	Cluster_1
jenis kelamin	2	2
status perkawinan	4	3
tingkat pendidikan	5	5
keterampilan	2	2
umur	23	38

Atribut yang berbeda antara *cluster* satu dengan yang lain terletak pada atribut status perkawinan dan umur. Seperti diperlihatkan pada Gambar 7 *cluster* 0 berjumlah 214 dan *cluster* 1 berjumlah 36 berarti bahwa para pencari kerja di

Dinas Tenaga Kerja dan Transmigrasi kab. Bantul didominasi oleh SMK dengan berjenis kelamin laki-laki dan memiliki keterampilan.

Cluster 0 memiliki jumlah data yang lebih besar karena rentang umur yang dominan berkisar antara 18 hingga 30 tahun, dengan sebagian besar data memiliki status perkawinan yang belum kawin atau belum menikah. Oleh karena itu, data yang memiliki umur 31 tahun ke atas dikategorikan sebagai bagian dari *Cluster 1*, dengan mayoritas data dalam *cluster* tersebut memiliki status perkawinan kawin atau sudah menikah.

4. Kesimpulan

Kesimpulan yang di dapat dari hasil penelitian yang telah dilakukan oleh peneliti pada kasus penerapan algoritma K-Means pada pengelompokan pencari kerja yang terdaftar di Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul, sebagai berikut :

Hasil *clustering K-Means* yang menunjukkan bahwa *Cluster_0* memiliki jumlah data yang lebih besar dengan 214 data didominasi rentang usia 18 – 30 tahun yang merupakan periode produktif untuk mencari pekerjaan. Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul dapat memberikan bimbingan karir untuk membantu pencari kerja menentukan jalur karir yang sesuai dengan minat, keterampilan, dan tujuan mereka. Dan *Cluster_1* memiliki data yang lebih sedikit dengan 36 data didominasi oleh rentang usia 31 – 56 tahun, Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul dapat menyediakan pelatihan lanjutan dan sertifikasi untuk meningkatkan daya saing. Dan hasil implementasi algoritma *K-Means* pada data pencari kerja yang telah diuji dengan $k=2$ sampai dengan $k=10$ menghasilkan *cluster* terbaik yaitu $k=2$ dengan nilai *Davies Bouldin Index* (DBI) sebesar 0.523, diantaranya *Cluster_0* terdiri dari 214 anggota, sementara *Cluster_1* terdiri dari 36 anggota. Oleh karena itu, semakin kecil nilai *Davies Bouldin Index* (DBI), semakin tinggi tingkat kemiripan data dalam satu kelompok.

Referensi

- [1] Ningsih, W., & Abdullah, F., "Analisis Perbedaan Pencari Kerja dan Lowongan Kerja Sebelum dan Pada Saat Pandemi Covid-19 di Kota Malang," *Journal of Regional Economics Indonesia*, vol. 2, no. 1, pp. 42-56, 2021.
- [2] Rohaniyah, S., & Purnamasari, A. I., "Clustering Data Pencari Kerja Menurut Tingkat Pendidikan Menggunakan Algoritma K-Means," *Jurnal Minfo Polgan*, vol. 12, no. 1, 2023.
- [3] P. Pa, A. M. H. Pardede, dan S. Rahmadani, "Pengelompokan Data Pencari Kerja Terdaftar Berdasarkan Umur dan Pendidikan Menggunakan Metode K-Means Clustering di Dinas Tenaga Kerja dan Perindustrian Perdagangan Kota Binjai," *Seminar Nasional Informatika (SENATIKA)*, 2022.
- [4] Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul, "Bidang PTKPKT. (n.d.). Data Pencari Kerja.," Dinas Tenaga Kerja dan Transmigrasi Kabupaten Bantul, [Online]. Available: disnakertrans.bantulkab.go.id/hal/lain-lain-bidang-ptkpkt-data-pencari-kerja. [Accessed 30 November 2023].
- [5] Agusta, Yudi, PhD, *K-Means Penerapan, Permasalahan dan Metode Terkait*, 2007.
- [6] Wicaksono, Y. , "Segmentasi Pelanggan Bisnis Dengan Multi Kriteria Menggunakan K-Means," *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 1, no. 2, pp. 45-53, 2019.
- [7] Wicaksono, Y., "SEGEMENTASI POTENSI DAERAH WISATA DI KOTA BIMA MENGGUNAKAN K-MEANS," *Indonesian Journal of Informatics and Research*, vol. 2, no. 2, pp. 26-32, 2021.
- [8] Butsianto, S., & Saepudin N., "Penerapan Data Mining Terhadap Minat Siswa Dalam Mata Pelajaran Matematika Dengan Metode K-Means," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 3, no. 1, pp. 51-59, 2020.
- [9] B. Bangoria, N. Mankad and V. Pambhar, "A Survey On Efficient Enhanced K-Means Clustering Algorithm," *International Journal For Scientific Research & Development*, no. 9, pp. 1698-1700, 2013.
- [10] Jollyta, D., Siddik, M., Mawengkang, H., & Efendi, S., *Teknik Evaluasi Cluster Solusi Menggunakan Python dan RapidMiner, Daerah Istimewa Yogyakarta: Deepublish*, 2021.
- [11] Xiao, J., Lu, J., & Li, X, "Davies Bouldin Index based hierarchical initialization Kmeans.," vol. 21, no. 6, p. 1327–1338, 2017.
- [12] Wicaksono, Y., Nendra Pratama, U., Nurhasanah, S., Utari Ramadania, T., & Juslan, W., "Student Admission Promotion Strategy Mapping Using K-Means," *Systemic: Information System and Informatics Journal*, vol. 7, no. 1, pp. 42-48, 2021.
- [13] C. Nas, "Data Mining Pengelompokan Bidang Keahlian Mahasiswa Menggunakan Algoritma K-Means (Studi Kasus : Universitas Cic Cirebon)," *Syntax: Jurnal Informatika*, vol. 09, no. 1, pp. 01-14, 2020.
- [14] Angra, S., & Ahuja, S., "Analysis of Student's Data Using Rapid Miner," *Journal on Today's Ideas - Tomorrow's Technologies*, vol. 4, pp. 49-58, 2016.
- [15] Hilda. S. D, Voutama, A., Umaidah, Y, "Analisis Daftar Pemilih Tetap Pemilihan Gubernur dan Wakil Gubernur Menggunakan Algoritma K-Means," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 10, no. 3, pp. 398-408, 2023.
- [16] F. Handayani, "Aplikasi Data Mining Menggunakan Algoritma K-Means," *Jurnal Teknologi dan Informasi (JATI)*, vol. 12, no. 1, pp. 46-63, 2022.