



Artikel

Penerapan Algoritme *K-Means* Dalam Mengelompokan Atlet PPLP Keolahragaan Di Jawa Barat

Fiqri Azzahra Sulaeman¹, Bani Nurhakim²

^{1,2}, STMIK IKMI, Manajemen Informatika, Cirebon, Indonesia

SUBMISSION TRACK

Recieved: Jan 22, 2023

Final Revision: March 12, 2023

Available Online: March 24, 2023

KEYWORD

Atlet PPLP, *K-Means*, *Clustering*

KORESPONDENSI

Phone: 089669880815

E-mail: fiqrisulaeman01@gmail.com

A B S T R A K

Pusat Pendidikan dan Latihan Pelajar (PPLP) adalah program dari Kementerian Pemuda dan Olahraga yang bertujuan untuk membina dan melatih atlet muda dalam berbagai cabang olahraga. PPLP telah ada di Jawa Barat sejak tahun 1992. Algoritma *K-Means* digunakan untuk mengatasi kesulitan dalam mengelompokkan wilayah berdasarkan jumlah peserta atlet PPLP setiap tahunnya dalam pencarian kelompok terbaik. Algoritme *K-Means* adalah metode pembelajaran *unsupervised* yang digunakan untuk mengelompokkan dataset yang belum terlabel ke dalam kelompok yang berbeda. Dalam penelitian ini, aplikasi *RapidMiner* digunakan dan data diambil dari Open Data Jawa Barat yang memiliki 243 data dari tahun 2013 hingga tahun 2021. Berdasarkan hasil analisis, teridentifikasi bahwa ada 5 kelompok atlet yang dibedakan berdasarkan jumlah atlet. Kelompok 0 tergolong *cluster* rendah, kelompok 1 tergolong *cluster* sedang, kelompok 2 tergolong *cluster* tinggi, kelompok 3 tergolong *cluster* sangat tinggi, dan kelompok 4 tergolong *cluster* sangat tinggi. Setiap kelompok memiliki jumlah atlet yang tidak sama dan memiliki wilayah yang tidak sama pula. Dalam kajian ini, 10 pengujian dilakukan dengan memanfaatkan beragam nilai K, dan hasil teroptimalkan teridentifikasi pada K sebanyak 5 dengan *Davies Bouldin Index* sebesar 0.958.

PENGANTAR

Dalam kehidupan sehari-hari, manusia menjalankan aktivitasnya dengan berbagai macam kegiatan. Tetapi disela-sela kesibukan, tubuh harus tetap terjaga dengan berolahraga walaupun hanya beberapa menit perhari. Olahraga merupakan proses yang terstruktur dari aktivitas, upaya, atau kegiatan

yang tepat untuk memperluas, membangun, dan meningkatkan potensi fisik dan spiritual [1]. Olahraga mempunyai peran seperti mesin pembentuk karakter dan pembangunan bangsa, karena misi olahraga adalah membangun jiwa raga dan mentalitas bangsa yang diperlukan untuk pembinaan atlet [2]. Dengan konsistensi dalam berolahraga, tidak sedikit orang yang menjadikan olahraga

sebagai ajang dalam meniti karir hingga menjadi atlet. Atlet merupakan olahragawan yang berjuang untuk menggapai prestasi maksimal [3]. Profesionalisme harus tetap dijaga, sehingga para atlet harus meningkatkan kualitas keterampilannya dan mampu meningkatkan prestasinya pada setiap kompetisi yang diikutinya. Kualitas pada kemampuan profesional seorang atlet meliputi konsentrasi yang tinggi, kompetensi profesional, kemampuan bertindak secara profesional dalam menghadapi masalah eksternal dan internal serta pengetahuan dan keterampilan yang terjaga [4]. Atlet yang bercita-cita meningkatkan prestasi mereka dalam cabang olahraga masing-masing dapat mengikuti Pusat Pendidikan dan Latihan Pelajar (PPLP). Dalam penelitian ini, terdapat data atlet PPLP sejumlah 243 data yang meliputi jumlah atlet dan wilayah mereka dari tahun 2013 hingga tahun 2021. Akan tetapi, terdapat masalah dalam mengelompokkan wilayah mana yang memiliki *cluster* optimum. Untuk mengatasi permasalahan ini, penulis menggunakan algoritme *clustering* dengan metode *K-Means*.

I. METODE

Metode pengumpulan informasi yang digunakan dalam kajian ini adalah melalui teknik pengumpulan data sekunder, yang didapatkan dari website Open Data Jabar. Informasi data tersebut tidak diambil melalui pendekatan langsung, namun melalui berbagai metode pengumpulan data, dan metode-metode tersebut meliputi :

1. Dataset mengenai atlet PPLP wilayah Jawa Barat pada tahun 2013 sampai tahun 2021 yang berasal dari *website* Open Data Jabar
2. Hasil dari dataset tersebut bisa dipergunakan untuk penelitian dalam bentuk excel kemudian disimpan kedalam aplikasi RapidMiner.

Tabel 1. Sampel Data Atlet PPLP di Jawa Barat

| Id | Nama_kota_kabupa tn | Jumlah_ atlet | tahun |
|-----------|--------------------------------|--------------------------|--------------|
| 1 | KABUPATEN BOGOR | 7 | 2013 |
| 2 | KABUPATEN SUKABUMI | 9 | 2013 |
| 3 | KABUPATEN CIANJUR | 5 | 2013 |
| 4 | KABUPATEN BANDUNG | 16 | 2013 |
| 5 | KABUPATEN GARUT | 1 | 2013 |
| 6 | KABUPATEN TASIKMALAYA | 3 | 2013 |
| 7 | KABUPATEN CIAMIS | 5 | 2013 |
| 8 | KABUPATEN KUNINGAN | 2 | 2013 |
| 9 | KABUPATEN CIREBON | 3 | 2013 |
| 10 | KABUPATEN MAJALENGKA | 4 | 2013 |
| ... | ... | ... | ... |
| 242 | KOTA TASIKMALAYA | 2 | 2021 |
| 243 | KOTA BANJAR | 1 | 2021 |

1.1. Data Mining

Data *mining* merupakan suatu tahapan yang bersifat berulang-ulang dan berkesinambungan untuk menemukan format dan rancangan baru yang berguna dan mudah dipahami dalam basis data yang sangat luas. Proses ini melibatkan penelusuran arus dan format yang diharapkan dalam basis data yang luas untuk membantu tahapan menentukan kebijakan di masa yang akan datang. Format-format ini dikenali oleh perangkat khusus yang mampu memberikan analisis data yang berguna dan berpendapat, dan mungkin menggunakan alat pendukung keputusan lain untuk dipelajari lebih lanjut [5].

1.2. Clustering

Clustering adalah teknik pembagian titik data dikelompokkan menjadi beberapa grup sehingga titik data dalam satu kelompok

memiliki kesamaan yang lebih besar daripada titik data dalam grup yang berbeda, hanya dengan memperhatikan data yang tersedia dalam setiap titik data [6]. *Clustering* dalam bahasa lain dapat didefinisikan sebagai proses pengelompokan data tanpa pengawasan, dimana teknik pembelajaran tak berlabel digunakan untuk membagi informasi ke dalam beberapa kelompok yang memiliki kesamaan tanpa membutuhkan fase pembelajaran terlebih dahulu [7]. Terdapat beberapa metode yang bisa digunakan untuk mengklasifikasikan data, termasuk teknik K-Means, K-Medoid, K-Mode, Hierarchical Clustering, dan lain sebagainya [8].

1.3. K-Means

Algoritma *K-Means* merupakan salah satu cara pengelompokan informasi yang berdasarkan pada setiap item data terkait dengan titik pusat kluster yang paling dekat. Prinsip dari *K-Means* adalah untuk membagikan informasi ke beberapa grup sehingga item data pada setiap kelompok memiliki kesamaan satu sama lain dan tidak mirip dengan item data pada kelompok lain [9]. Langkah-langkah melakukan proses *clustering* menggunakan metode *K-Means* adalah sebagai berikut [10]:

1. Menentukan jumlah kluster yang diinginkan dengan memilih nilai k .
2. Inisialisasi centroid dapat dilakukan dengan berbagai cara, biasanya dengan memilih secara acak dari data yang ada.
3. Untuk menentukan kelompok setiap item data, jarak antara setiap data input dan *centroid* dihitung menggunakan rumus jarak *Euclidean*. Proses ini bertujuan untuk menemukan jarak terdekat dari setiap data dengan *centroid*. Berikut adalah rumus dari *Euclidean Distance* :

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

dimana :

De merupakan *Euclidean Distance*,
 i merupakan banyaknya objek,
 (x, y) yaitu koordinat objek,
 (s, t) yaitu koordinat *centroid*

4. Menggolongkan setiap data berdasarkan jarak terdekatnya dengan *centroid*.
5. memperbarui nilai *centroid* dengan menggunakan nilai rata-rata dari informasi yang ada dalam setiap *cluster* :

$$v_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

dimana :

v_{ij} merupakan *centroid*/rata-rata *cluster* ke- i untuk variable ke- j ,
 N_i yaitu jumlah data yang menjadi anggota *cluster* ke- i ,
 i, k yaitu indeks dari *cluster*
 j merupakan indeks dari variabel
 x_{kj} merupakan nilai data ke- k yang ada di dalam *cluster* tersebut untuk variabel ke- j

6. Melakukan proses pengulangan dari langkah 2 sampai 5 sampai elemen dalam setiap grup tidak berubah.

1.4. Davies Bouldin Index

Metode *Davies Bouldin Index* digunakan untuk menilai kevalidan jumlah *cluster* optimal pada suatu teknik pengelompokan. DBI mengukur kohesi *cluster* dengan menentukan jumlah jarak data terhadap titik pusat dari cluster yang bersangkutan [11]. Prinsipnya adalah meminimalkan jarak intra-*cluster* dan memaksimalkan jarak antar-*cluster*. Semakin baik skema kluster ditunjukkan oleh nilai DBI yang lebih rendah [12]. Berikut adalah langkah-langkah untuk menghitung DBI [13]:

$$DBI = \frac{1}{k} \sum_{a=1}^k R_a$$

dengan

$$R_a = \max_{b = 1, \dots, k, a \neq b} R_{ab} ,$$

$$R_{ab} = \frac{s_a + s_b}{d(V_a, V_b)}$$

dimana :

k = Jumlah kluster

R_{ab} = Ukuran kemiripan antara kluster ke- a dan kluster ke- b

s_a = Ukuran dispersi kluster ke- a , $a = 1, 2, \dots, k$

$$s_a = \left[\frac{1}{n_a} \sum_{T_i \in C_a, i=1}^{n_a} d^2(T_i, V_a) \right]^{\frac{1}{2}} , d^2(T_i, V_a) = (d(T_i, V_a))^2$$

dimana :

n_a = Banyaknya anggota kluster ke- a , $a = 1, 2, \dots, k$

C_a = Kluster ke- a

T_i = Anggota ke- i pada kluster ke- a , $a = 1, 2, \dots, k$

Untuk mengukur jarak antara anggota ke- i dari kluster ke- a (T_i) dan *centroid* kluster ke- a (V_a), DBI mempergunakan ukuran sederhana ketidakcocokan untuk menilai kualitas skema kluster dikenal sebagai jarak $d(T_i, V_a)$

$$d(T_i, V_a) = \sum_{j=1}^m \delta(x_{ij}, v_{aj})$$

dengan

$$\delta(x_{ij}, v_{aj}) = \begin{cases} 0 & x_{ij} = v_{aj} \\ 1 & x_{ij} \neq v_{aj} \end{cases}$$

dimana :

x_{ij} = nilai dari variabel ke- j pada T ke- i

v_{aj} = nilai ke- j pada *centroid* kluster ke- a

m = Jumlah variabel

Menghitung jarak antara *centroid* kluster ke- a (V_a) dan *centroid* kluster ke- b (V_b) dengan menggunakan ukuran ketidaksamaan pencocokan sederhana adalah langkah sebagai berikut:

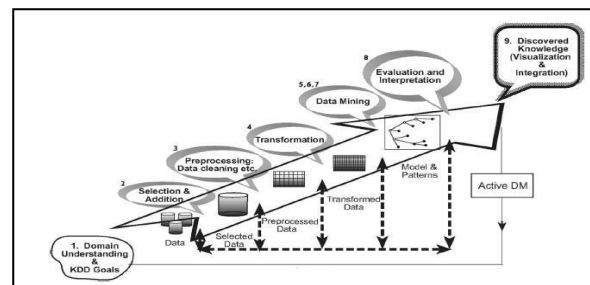
$$d(V_a, V_b) = \sum_{j=1}^m \delta(V_{aj}, V_{bj})$$

dengan

$$\delta(V_{aj}, V_{bj}) = \begin{cases} 0 & V_{aj} = V_{bj} \\ 1 & V_{aj} \neq V_{bj} \end{cases}$$

dimana nilai ke- j pada *centroid* kluster ke- b , dikenal sebagai (V_{bj}).

1.5. KDD



Gambar 1. Tahapan penelitian [14].

Knowledge Discovery in Database (KDD) merupakan suatu proses yang memanfaatkan teknik pemrosesan data untuk memperoleh informasi yang sesuai dengan spesifikasi ukuran dan batas dengan memanfaatkan basis data melalui proses pembersihan dan pemrosesan data, pengambilan sampel, dan transformasi dari basis data [15]. Berikut merupakan tahapan – tahapan dan proses yang ada pada Knowledge Discovery in Database :

1. Selection

Di awal tahapan terdapat tahap selection yang merupakan proses analisis data yang berkaitan dari database karena biasanya tidak semua data memerlukan proses pembuatan kebijakan dalam data mining. Dalam hal data atlet PPLP di Jawa Barat dari tahun 2013 hingga 2021, sejumlah

243 data dipilih melalui proses seleksi.

2. *Pre-processing*

Tahap berikutnya adalah *pre-processing*, di mana data yang sudah dikumpulkan akan dibersihkan dari kesalahan atau ketidaklengkapan data. Proses ini meliputi tahap seperti data yang kosong, ganda ataupun *missing* maka dilakukan proses *cleaning* agar sesuai dengan format yang diinginkan.

3. *Transformation*

Pada tahap *transformation*, data akan dikonversi dari tipe atribut kategori ke numerik. Contohnya, nama kabupaten/kota akan dikonversi menjadi numerik agar dapat diproses dengan algoritme *K-Means* pada data *mining*.

4. *Data mining*

Tahap utama dari proses data *mining* merupakan data *mining* itu sendiri. Jumlah *cluster* harus ditentukan sebelumnya untuk mencapai nilai *K* yang optimal. Proses ini menggunakan *RapidMiner* sebagai *tool*.

5. *Interpretation / Evaluation*

Pada proses ini, data yang sudah diproses akan diukur menggunakan operator kinerja untuk menghitung *Davies Bouldin Index (DBI)*, dengan tujuan untuk memperoleh nilai *K* yang optimal.

II. HASIL

2.1 Selection

Pada tahap awal yaitu *selection* yang bertujuan untuk memisahkan data yang digunakan dan yang tidak digunakan.

| Row No. | nama_kabupaten_kota | jumlah_atlet | tahun |
|---------|-----------------------|--------------|-------|
| 1 | KABUPATEN BOGOR | 7 | 2013 |
| 2 | KABUPATEN SUKABUMI | 9 | 2013 |
| 3 | KABUPATEN CIANJUR | 5 | 2013 |
| 4 | KABUPATEN BANDUNG | 16 | 2013 |
| 5 | KABUPATEN GARUT | 1 | 2013 |
| 6 | KABUPATEN TASIKMALAYA | 3 | 2013 |
| 7 | KABUPATEN CIAMIS | 5 | 2013 |
| 8 | KABUPATEN KUNINGAN | 2 | 2013 |
| 9 | KABUPATEN CIREBON | 3 | 2013 |
| 10 | KABUPATEN MAJALENGA | 4 | 2013 |
| 11 | KABUPATEN SUMEDANG | 8 | 2013 |
| 12 | KABUPATEN INDRAMAYU | 5 | 2013 |
| 13 | KABUPATEN SUBANG | 5 | 2013 |
| 14 | KABUPATEN PURWAKARTA | 5 | 2013 |
| 15 | KABUPATEN KARAWANG | 6 | 2013 |

Gambar 2. Data yang akan digunakan

Dari semua 8 karakteristik, yaitu id, kode provinsi, nama provinsi, kode kabupaten/kota, nama kabupaten/kota, jumlah atlet, satuan, dan tahun, hanya tiga diantaranya yang dipilih sebagai karakteristik utama, yaitu nama kabupaten/kota, jumlah atlet, dan tahun.

2.2 Pre-Processing

Tahap *Pre-Processing/Cleaning* melakukan proses pengaturan dan pembersihan data sebelum dilakukan analisis atau pemodelan. Ini bertujuan untuk memverifikasi bahwa data yang akan dipakai sudah siap untuk diproses yang memiliki kualitas baik, akurat dan sesuai dengan kebutuhan. Langkah-langkah ini meliputi penanganan data yang kosong, ganda atau *missing*.

Gambar 3. Hasil data *cleaning*

Setelah dilakukan pembersihan data yang menghasilkan tidak adanya perubahan data dan masih tetap sama berjumlah 243 data dari tahun 2013 sampai tahun 2021, dikarenakan data yang digunakan merupakan data lengkap

tidak ada data yang kosong, ganda ataupun *missing*.

2.3 Transformation

Dalam tahap transformasi data, atribut dipilih dan ditransformasikan ke dalam bentuk yang sesuai dengan algoritme *K-Means*, yang memiliki dua tahap. Tahap pertama untuk mengkonversi atribut non-numerik menjadi representasi numerik, operator *Nominal to Numerical* digunakan agar dapat diterima oleh algoritme *K-Means*.

Gambar 4. Hasil transformasi data

Tahapan yang kedua yaitu menggunakan operator *Normalize* untuk melakukan normalisasi. Operator ini berfungsi sebagai memperkecil jarak jumlah atlet antara nilai 0.0 hingga 1.0. Operator ini dilakukan karena besarnya selisih jarak antar jumlah atlet yang akan menyebabkan tidak optimalnya pengelompokan ini.

III. PEMBAHASAN

Perlu adanya visualisasi yang efektif untuk menunjukkan hasil dari proses data *mining* sehingga mudah dipahami oleh pihak terkait.

Gambar 5. Hasil normalisasi

2.4 Data Mining

Tahap selanjutnya memasukkan operator *Clustering* yang berfungsi untuk menemukan *cluster* atau pengelompokan data yang memiliki karakteristik tertentu. Dalam penelitian ini, digunakan algoritme *K-Means* untuk melakukan proses *clustering* dan menggunakan 10 *cluster* max perulangan 10 kali putaran dan jenis *measure types* nya adalah *NumericalMeasures* dengan menggunakan metode perhitungan *EuclideanDistance* dan *max optimization steps* nya berjumlah 100.

Gambar 6. Hasil clustering

Tabel 2. Jumlah anggota atlet pplp berdasarkan cluster

| Cluster | Jumlah Atlet | Tingkat penyumbang |
|-----------|--------------|--------------------|
| Cluster 0 | 190 | Rendah |
| Cluster 1 | 203 | Sedang |
| Cluster 2 | 277 | Tinggi |
| Cluster 3 | 138 | Sangat rendah |
| Cluster 4 | 396 | Sangat tinggi |

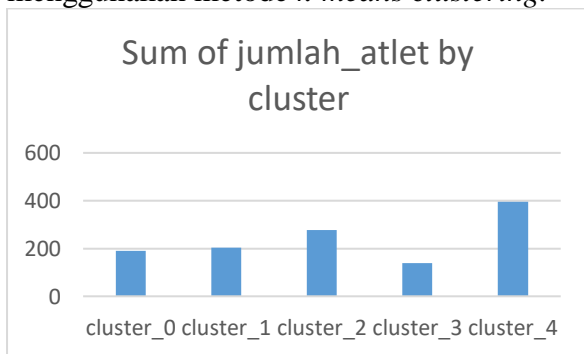
Untuk mendapatkan nilai K yang paling optimum, dilakukan 10 kali percobaan. Nilai K yang paling optimum adalah nilai K dengan

hasil DBI mendekati 0. Dari 10 kali percobaan, nilai K yang paling optimum terdapat pada K = 5 dengan hasil DBI nya adalah 0.958.

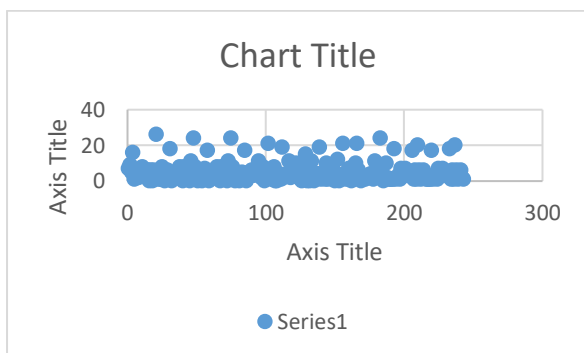
Tabel 3. Hasil DBI untuk menentukan k optimum

| No | Kelompok (K) | Davies Bouldin Index |
|----|--------------|----------------------|
| 1 | k2 | 1.285 |
| 2 | k3 | 1.057 |
| 3 | k4 | 1.033 |
| 4 | k5 | 0.958 |
| 5 | k6 | 0.996 |
| 6 | k7 | 0.963 |
| 7 | k8 | 0.998 |
| 8 | k9 | 0.966 |
| 9 | k10 | 0.978 |

Hasil pengelompokan data atlet PPLP Keolahragaan di Jawa Barat dapat ditemukan setelah dilakukan proses *clustering* dengan menggunakan metode *k-means clustering*.



Gambar 7. Hasil cluster jumlah data atlet



Gambar 8 Scatter Plot persebaran data atlet PPLP di Jawa Barat

Dalam hal ini, *scatter plot* menampilkan persebaran atlet PPLP keolahragaan di Jawa

Barat melalui metode *k-means clustering* akan menampilkan persebaran atlet tersebut berdasarkan dua variabel yang dipilih, yaitu jumlah atlet dan *cluster* wilayah. Kemudian, metode *k-means clustering* akan digunakan untuk mengelompokkan titik-titik atlet ke dalam beberapa kelompok (*cluster*) berdasarkan jarak antar titik. Hasil dari *scatter plot* ini akan memperlihatkan bagaimana persebaran atlet PPLP keolahragaan di Jawa Barat dibedakan ke dalam beberapa kelompok berdasarkan jumlah atlet dan *cluster* wilayah. Ini akan membantu kita untuk mengetahui pola persebaran atlet dan memberikan wawasan lebih lanjut mengenai tingkat prestasi dan jumlah atlet PPLP keolahragaan di Jawa Barat.

IV. KESIMPULAN DAN SARAN

4.1 Kesimpulan

1. Berdasarkan hasil pengelompokan atlet PPLP Keolahragaan di wilayah Jawa Barat selama periode 2013 hingga 2021, terdapat lima kelompok yang ditemukan. Masing-masing kelompok memiliki jumlah atlet dan wilayah yang berbeda-beda, dengan tingkat prestasi yang berbeda-beda pula. Kelompok 0, 1, 2, 3, dan 4 termasuk dalam kategori rendah, sedang, tinggi, paling rendah, dan paling tinggi. Hasil ini dapat membantu untuk mengetahui pola persebaran atlet dan memperlihatkan tingkat prestasi atlet PPLP Keolahragaan di wilayah Jawa Barat selama periode tersebut.
2. Penelitian ini melakukan 10 kali percobaan dengan K yang berbeda. Dan diketahui dari 10 kali percobaan, nilai K yang paling optimum terdapat pada K = 5 dengan hasil DBI nya adalah 0.958.

4.2 Saran

Pada penelitian berikutnya ada beberapa rekomendasi yang bisa disampaikan, yaitu sebagai berikut :

1. Merekomendasikan Pengelompokan algoritme terbaik menggunakan Algoritme *K-Means*,
2. perbanyak percobaan terhadap algoritme selain *K-Means* yang digunakan karena setiap algoritme unik dan mempunyai karakteristik yang berbeda-beda, dan
3. Harapannya, hasil penelitian ini akan memberikan dasar yang kuat untuk penelitian selanjutnya untuk mengelompokkan jumlah atlet berdasar wilayah di Jawa Barat.

REFERENCES

- [1] N. A. Wahyudi, "Peran perkembangan industri olahraga dan rekreasi dalam menumbuhkan jiwa kewirausahaan bagi generasi muda," *Pros. SNIKU (Seminar Nas. Ilmu Keolahragaan UNIPMA)*, vol. 1, no. 1, pp. 34–42, 2018.
- [2] R. M. Aguss and R. Yuliandra, "Persepsi Atlet Futsal Putra Universitas Teknokrat Indonesia Terhadap Hipnoterapi Dalam Meningkatkan Konsentrasi Saat Bertanding," *J. Penjaskesrek*, vol. 7, no. 2, pp. 274–288, 2020.
- [3] I. G. Putu, N. Adi, A. Agung, and O. Perdana, "Analisis Kondisi Fisik Atlet di Masa Pandemi Covid 19," *Pros. Semin. Nas. IPTEK Olahraga*, pp. 1–5, 2021.
- [4] K. A. Wati and M. Jannah, "Hubunga Antara Kejenuhan dengan Motivasi Berprestasi Pada Atlet Sepak Bola," *J. Psikol.*, vol. 08, no. 03, pp. 126–136, 2021, [Online]. Available: <https://ejournal.unesa.ac.id/index.php/character/article/view/41205>
- [5] E. D. Sikumbang, "Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori," *J. Tek. Komput. AMIK BSI*, vol. Vol 4, No., no. September, pp. 1–4, 2018.
- [6] V. Herlinda and D. Darwis, "Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means," *Darwis, Dartono*, vol. 2, no. 2, pp. 94–99, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTISI>
- [7] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 1, p. 62, 2019, doi: 10.26418/jp.v5i1.29611.
- [8] D. Ayu, I. Cahya, D. Ayu, and K. Pramita, "Analisis Perbandingan Metode Elbow dan Sillhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali," vol. 9, no. 3, 2019.
- [9] Y. Prayoga, H. S. Tambunan, and I. Parlina, "Penerapan Clustering Pada Laju Inflasi Kota Di Indonesia Dengan Algoritma K-Means," *BRAHMANA J. Penerapan Kecerdasan Buatan*, vol. 1, no. 1, pp. 24–30, 2019, doi: 10.30645/brahmana.v1i1.4.
- [10] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *J. Tekno Kompak*, vol. 15, no. 2, p. 25, 2021, doi: 10.33365/jtk.v15i2.1162.
- [11] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *J. Sains dan Manaj.*, vol. 9, no. 1, pp. 95–100, 2021, [Online]. Available: <https://ejournal.bsi.ac.id/ejurnal/index.php/evolusi/article/view/10428/4839>
- [12] W. Gie *et al.*, "Perbandingan Euclidean dan Manhattan Untuk Optimasi Cluster Menggunakan Davies Bouldin Index : Status Covid-19 Wilayah Riau," vol. 2, no. April, pp. 187–191, 2020.
- [13] A. Badruttamam, S. Sudarno, and D. A. I. Maruddani, "PENERAPAN ANALISIS KLASER K-MODES DENGAN VALIDASI DAVIES BOULDIN INDEX DALAM MENENTUKAN KARAKTERISTIK KANAL YOUTUBE DI INDONESIA (Studi Kasus: 250 Kanal YouTube Indonesia Teratas Menurut Socialblade)," *J. Gaussian*, vol. 9, no. 3, pp. 263–272, 2020, doi: 10.14710/j.gauss.v9i3.28907.
- [14] M. R. Muttaqin and M. Defriani, "Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa," *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.
- [15] Q. Widayati, "Penerapan Data Mining Menggunakan Metode Teknik Classification Untuk Melihat Potensi Kepatuhan Wajib Pajak Bumi Dan Bangunan," *J. Ilm. Matrik*, vol. 20, no. 2, pp. 157–168, 2019, doi: 10.33557/jurnalmatrik.v20i2.119.

BIOGRAPHY

Fiqri Azzahra Sulaeman, lahir di Cirebon pada tanggal 29 Agustus 2001. Sekolah Dasar dilaksanakan di SD Kebon Pelok, SMP 7 Cirebon dan SMA 6 Cirebon. Menyelesaikan pendidikan Diploma 3 (D3) Program Studi Manajemen Informatika pada tahun 2023 di STMIK IKMI Cirebon.

Bani Nurhakim, menjabat sebagai dosen tetap di STMIK IKMI Cirebon.