

Implementasi Metode DBSCAN Untuk Pengelompokkan Sebaran Wilayah Asal Mahasiswa Sebagai Strategi Promosi Kampus

Andre Gio Pane

Fakultas Ilmu Komputer dan Teknologi Informasi, Prodi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: agiopane@gmail.com

Abstrak—Teknologi database saat ini memungkinkan untuk menyimpan sejumlah data dalam jumlah yang sangat besar dan terakumulasi namun disinilah awal timbulnya persoalan ledakan data, contohnya pada Universitas. Universitas Budi Darma Medan masih melakukan analisis data secara manual untuk menentukan rencana strategi promosi dari sebaran mahasiswa berdasarkan geografis. Sebaran mahasiswa dapat dilakukan dengan cara mengelompokkan data mahasiswa berdasarkan kesamaan ciri dari data menggunakan metode clustering dengan algoritma dbscan. Pembuatan aplikasi data mining bertujuan untuk memudahkan menganalisis pengelompokan sebaran mahasiswa. Data yang digunakan adalah data angkatan mahasiswa tahun 2018 sampai tahun 2019. Pada penelitian ini aplikasi dibangun dengan menggunakan Rapid Miner. Teknik ini dapat menentukan cluster dari bentuk data yang tidak beraturan dan dapat menangani noise secara efektif. Penelitian ini berfokus pada pengimplementasian Metode DBSCAN pada proses Pengelompokkan sebaran asal wilayah mahasiswa sebagai strategi promosi kampus Universitas Budi Darma Medan untuk kedepannya. Di harapkan sistem yang akan dibangun sesuai dengan harapan peneliti, sehingga dapat mengelompokkan sebaran asal wilayah mahasiswa sebagai strategi promosi kampus Universitas Budi Darma Medan

Kata Kunci: Data Mining; Clustering; DBSCAN; Mahasiswa; Promosi

Abstract—Current database technology makes it possible to store huge amounts of accumulated data but this is where the data explosion problem arises, for example at the University. Budi Darma University Medan is still analyzing data manually to determine the promotion strategy plan of the geographic distribution of students. The distribution of students can be done by grouping student data based on the similarity of the characteristics of the data using the clustering method with the dbscan algorithm. Making data mining applications aims to make it easier to analyze student distribution groupings. The data used is data from the 2018 to 2019 batch of students. In this study the application was built using Rapid Miner. This technique can determine clusters of irregular data shapes and can handle noise effectively. This research focuses on the implementation of the DBSCAN method in the process of grouping the distribution of student areas from the university as a promotion strategy for the Budi Darma University campus in the future. It is hoped that the system will be built in accordance with the expectations of researchers, so that it can classify the distribution of student areas from as a Budi Darma University campus promotion strategy.

Keywords: Data Mining; Clustering; DBSCAN; Student; Promotion

1. PENDAHULUAN

Persaingan antarperguruan tinggi dalam menarik mahasiswa baru semakin ketat, sehingga diperlukan strategi promosi yang tepat sasaran dan efisien[1]. Universitas Budi Darma menghadapi permasalahan dalam menentukan lokasi promosi yang efektif, karena mahasiswa berasal dari berbagai daerah dengan distribusi yang beragam. Tanpa pendekatan analitis berbasis data, promosi sering kali membutuhkan biaya besar tetapi tidak selalu menghasilkan jumlah pendaftar yang optimal. Selain itu, pendekatan promosi konvensional masih dilakukan secara general tanpa mempertimbangkan faktor kepadatan sebaran mahasiswa, sehingga peluang potensial dari daerah tertentu sering terlewatkan. Hal ini mengakibatkan universitas kesulitan dalam merancang strategi yang berbasis pada data empiris yang valid. Oleh karena itu, diperlukan metode data mining yang mampu mengelompokkan wilayah asal mahasiswa secara akurat sebagai dasar penyusunan strategi promosi universitas.

Data mining merupakan pendekatan analitis yang mampu menggali pengetahuan baru dari basis data besar, termasuk dalam bidang pendidikan untuk mendukung strategi promosi kampus[2]. Salah satu metode populer dalam data mining adalah clustering, yang digunakan untuk mengelompokkan data berdasarkan kemiripan tertentu, sehingga dapat mengidentifikasi pola tersembunyi yang sebelumnya sulit dikenali[3]. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) termasuk metode clustering berbasis kepadatan yang efektif dalam menemukan kelompok data dengan bentuk tidak beraturan dan tahan terhadap data noise[4].

DBSCAN bekerja dengan mengidentifikasi titik inti berdasarkan jumlah tetangga dalam radius tertentu, kemudian mengembangkan cluster dengan menghubungkan titik-titik yang memiliki kepadatan serupa[2]. Berbeda dengan metode K-Means yang membutuhkan jumlah cluster di awal, DBSCAN tidak memerlukan jumlah cluster sebagai parameter, sehingga lebih fleksibel untuk data pendidikan yang heterogen [5]. Fleksibilitas ini menjadikan DBSCAN sangat relevan untuk mengelompokkan sebaran asal mahasiswa yang bervariasi, baik dari daerah perkotaan maupun pedesaan, sehingga universitas dapat mengidentifikasi potensi daerah dengan lebih tepat[6].

Dalam konteks strategi promosi kampus, penggunaan DBSCAN dapat membantu universitas dalam memahami pola sebaran geografis mahasiswa berdasarkan asal sekolah atau daerah domisili. Informasi ini sangat penting untuk menentukan daerah mana yang lebih potensial dalam memberikan kontribusi mahasiswa baru, sehingga dapat meminimalkan biaya promosi yang besar namun tidak efektif [3]. Dengan demikian, penerapan DBSCAN dalam promosi pendidikan tinggi tidak hanya bersifat teknis, tetapi juga strategis karena mendukung pengambilan keputusan berbasis data[6].

Beberapa penelitian sebelumnya menunjukkan keberhasilan DBSCAN dalam berbagai bidang analisis data pendidikan maupun spasial, namun belum banyak yang mengaitkannya langsung dengan strategi promosi perguruan tinggi. Penelitian yang dilakukan Yanto, dkk pada tahun 2024 mengenai pemetaan pusat industri dengan algoritma DBSCAN menunjukkan bahwa pendekatan ini efektif dalam mengidentifikasi pusat-pusat industri berdasarkan distribusi spasial dan kepadatan data. Studi kasus di Kabupaten Situbondo menggunakan data industri kecil dan menengah (IKM) dari Dinas Koperasi, Perindustrian, dan Perdagangan setempat. Hasil analisis berhasil mengelompokkan lokasi industri menjadi beberapa kluster kepadatan tinggi yang ditetapkan sebagai pusat industri utama, sekaligus mengidentifikasi wilayah potensial untuk pengembangan pusat industri baru. Temuan ini memberikan dasar strategis bagi pemerintah daerah dalam mendukung perencanaan pembangunan wilayah dan ekonomi berkelanjutan[7].

Penelitian Andri dan Eko Riswanto (2023), penelitian ini berhasil mengelompokkan profil mahasiswa STMIK EL Rahma Yogyakarta berdasarkan IPK dan frekuensi kunjungan perpustakaan menggunakan algoritma DBSCAN. Hasil analisis pada 75 sampel mahasiswa dengan parameter epsilon 0,4 dan MinPts 2 menghasilkan 7 cluster yang distinct dengan 5 data noise, menunjukkan bahwa metode ini efektif dalam memetakan kebiasaan belajar mahasiswa serta memberikan gambaran nyata mengenai minat baca dan kaitannya dengan kinerja akademik[8].

Penelitian sebelumnya Richardo Anggara, dkk tahun 2022 mengenai Chronic Kidney Disease (CKD) memanfaatkan metode Density-Based Spatial Clustering of Applications with Noise (DBSCAN) untuk mengklasifikasi data pasien yang terdiagnosis CKD. Dataset yang digunakan terdiri dari 400 data medis dengan 24 atribut. Hasil penelitian menunjukkan bahwa penerapan DBSCAN dengan parameter epsilon 3.5 dan minimum sample 5 menghasilkan 2 kluster dengan nilai Silhouette 0.158, meskipun juga ditemukan cukup banyak noise.[9]

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis[10][11]. Definisi lain diantaranya adalah pembelajaran berbasis induksi (induction-based learning) yang merupakan proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari[12]. Data mining digunakan secara luas dalam berbagai bidang seperti pendidikan, pemasaran, dan kesehatan untuk menemukan pola tersembunyi yang berguna untuk pengambilan keputusan strategis.

2.2 Metode DBSCAN

DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) adalah algoritma *clustering* berbasis kepadatan yang mampu membentuk kluster dengan bentuk arbitrer serta mengidentifikasi *noise*. Metode ini bekerja dengan dua parameter utama, yaitu Eps (radius tetangga) dan MinPts (jumlah minimum titik dalam radius tersebut). Penelitian terbaru menunjukkan bahwa DBSCAN tetap relevan untuk data berukuran besar dan kompleks, namun berbagai pengembangan dilakukan untuk mengatasi keterbatasannya, seperti sensitivitas terhadap parameter dan efisiensi komputasi. Misalnya, DG-DBSCAN (Dual Grid) memperbaiki pemilihan parameter dan mengurangi waktu komputasi, sementara DBSCAN berbasis *block-diagonal graph* meningkatkan akurasi pada data berdimensi tinggi. Selain itu, survei terkini menekankan pentingnya varian DBSCAN dalam pemrosesan *big data* karena fleksibilitasnya pada dataset dengan kepadatan berbeda[13][14].

2.3 Promosi

Promosi pada hakekatnya adalah suatu komunikasi pemasaran, artinya aktivitas pemasaran yang berusaha menyebarkan informasi, mempengaruhi/membujuk, dan atau mengingatkan pasar sasaran atas perusahaan dan produknya agar bersedia menerima, membeli, dan loyal pada produk yang ditawarkan perusahaan. Promosi juga dapat didefinisikan sebagai suatu upaya perusahaan dalam mempengaruhi konsumen aktual maupun konsumen potensial agar mereka mau melakukan pembelian terhadap produk yang ditawarkan saat ini maupun di masa mendatang [15]. Dalam konteks perguruan tinggi, promosi berfungsi untuk memperkenalkan institusi kepada masyarakat luas sekaligus meningkatkan jumlah mahasiswa baru[16].

3. HASIL DAN PEMBAHASAN

Mahasiswa Universitas Budi Darma Medan banyak dari berbagai wilayah medan sampai dari luar daerah medan, karna ada beberapa fakultas yang siap untuk bersaing di era modern pada masa yang akan datang. Pengelompokkan sebaran wilayah ini dilakukan berdasarkan Wilayah asal Mahasiswa. Berdasarkan hasil pengamatan pengelompokkan wilayah di Universitas Budi Darma Medan ini diketahui belum maksimal dan efisien maka perlu dilakukan penelitian Pengelompokkan asal sebaran wilayah para Mahasiswa yang akan menjadi strategi promosi kampus Universitas Budi Darma dengan menggunakan metode DBSCAN. DBSCAN atau yang biasa dikenal

dengan sebutan *density based clustering* adalah salah satu contoh pelopor perkembangan teknik pengelompokan berdasarkan kepadatan.

Data Utama pembahasan dalam penelitian adalah data-data asal sebatan wilayah Mahasiswa yang akan menjadi strategi promosi kampus, dalam penelitian ini hanya data Mahasiswa Budi Darma Medan Pada Tahun 2018-2019 saja yang akan di olah dengan algoritma *DBSCAN*, sebagai referensi dapat di lihat pada tabel 1.

Tabel 1. Hasil Data Preprosesing

No.	Nama	Asal Sekolah	Asal Wilayah
1	Cici Alfiani Pradika Dita	2	1
2	Helena Wisarani Nazara	3	2
3	Yuni Efridayanti	2	1
4	HELMI FACHRI	3	3
5	Arby Hartama	5	1
6	Rama Aditya	4	1
7	Desi Rahmadani	1	3
8	Sella Novita Rizky	3	4
9	Manila Mendrofa	1	2
10	Andi Komang	1	4
11	Anissa Puji Arni	1	3
12	Ria Ramadhani Syahputri	2	4
13	Alfianus Harefa	2	2
14	Ahmad Ryandi Rafly Hutabarat	1	1
15	Andre Martin Pandiangan	4	1
16	Dessy Dhasmara Sukma	1	4
17	Dinda Meutia	2	4
18	Pola Pinterman Gea	2	2
19	Yantiria Halawa	1	2
20	Khoirunnas	1	1

Tabel 1 menampilkan hasil data preprocessing yang berisi daftar 20 nama siswa beserta asal sekolah (dalam bentuk kode angka 1–5) dan asal wilayah (kode angka 1–4). Dari tabel ini terlihat bahwa siswa berasal dari berbagai sekolah dan wilayah yang berbeda, misalnya ada yang berasal dari sekolah kode 1 hingga 5, serta wilayah kode 1 hingga 4. Pola distribusi menunjukkan keragaman asal siswa, di mana beberapa sekolah dan wilayah lebih sering muncul (misalnya sekolah 1 dan wilayah 1 cukup dominan), sementara yang lain lebih jarang. Hal ini menunjukkan bahwa data sudah berhasil disederhanakan ke dalam bentuk numerik sehingga lebih siap untuk dianalisis lebih lanjut.

3.1 Penerapan Metode DBSCAN

Metode clustering Density-Based Spatial Clustering of Application with Noise (DBSCAN) membangun region berdasarkan kepadatan koneksi. Algoritma yang dirancang oleh Esther dkk. Pada tahun 1996, DBSCAN mampu mengidentifikasi cluster kombinasi data spasial yang besar dengan melihat kepadatan lokal elemen database hanya dengan menggunakan satu parameter input. DBSCAN juga dapat menentukan apakah informasi tersebut telah diklasifikasikan.

Algoritma DBSCAN adalah bahwa untuk setiap titik dari sebuah *cluster*, *neighborhood* dari radius yang diberikan harus mengandung setidaknya jumlah minimum poin, yaitu, kepadatan *neighborhood* harus melebihi beberapa *threshold* ditetapkan. Algoritma ini membutuhkan dua parameter masukan untuk bisa melakukan clustering yaitu:

1. *Eps*, radius yang menentukan batas daerah *neighborhood* dari titik (*Eps-neighborhood*).
2. *MinPts*, jumlah minimum poin yang harus ada di *Eps- neighborhood* -. Adapun urutan algoritma dari DBSCAN secara umum memiliki 6 langkah yaitu:
 - a. *Arbitrary select a point p* (Pilih point p awal secara acak)
 - b. *Retrieve all points density-reachable from p wrt Eps and MinPts* (Ambil semua point yang *density reachable* terhadap titik p)
 - c. *If p is a core point, a cluster is formed* (Jika p adalah *core point* maka cluster terbentuk)
 - d. *If p is a border point, no points are density-reachable from p and DBSCAN visits the next point of the database* (Jika p adalah border point, tidak ada yang merupakan hubungan *density-reachable* dari p dan DBSCAN akan mengunjungi point selanjutnya dari database)
 - e. *Continue the process until all of the points have been processed* (Lanjutkan proses sampai semua point telah diproses).
 - f. *Result is independent of the order of processing the points* (Hasil yang didapatkan tidak terhantung dari urutan dari proses point yang diambil).

Dimana nantinya data akan diolah dengan menggunakan Metoda *desity-based* yaitu DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*). Dengan parameter input

- a. *MinPts* :3
- b. *Eps* : 2

Untuk menghitung jarak titik *core point* dengan *point* yang lain pada DBSCAN bisa dengan menggunakan rumus jarak biasa yaitu *Euclidean Distance*. Sesuai dengan algoritma DBSCAN langkah pertama adalah memilih titik pusat awal secara acak sesuai dengan keinginan. Contoh disini akan memakai titik B sebagai titik pusat awal untuk iterasi 1.

Iterasi I

Iterasi I adalah dengan mengitung jarak masing masing *point* atau titik terhadap titik pusat B. Berikut adalah perhitungan jarak masing-masing titik dengan *core point* untuk iterasi pertama, pada soal diatas dengan menggunakan rumus jarak diatas:

- a. *MinPts* :3
- b. *Eps* :2
- c. Pusat titik B

$$x_p = 3 ;$$

$$y_p = 2 \text{ Perhitungannya}$$

$$AB = \sqrt{(2 - 3)^2 + (1 - 2)^2} = 2$$

Dilakukan hal yang sama untuk titik yang lain, sehingga didapatkan perhitungan jarak untuk iterasi I sebagai berikut

Tabel 2. Hasil Jarak Iterasi I

Jarak	Hasil
AB	1,414214
BB	0
CB	1,414214
DB	1
EB	2,236068
FB	1,414214
GB	2,236068
HB	2
IB	2
JB	2,828427
KB	2,236068
LB	2,236068
MB	1
NB	2,236068
OB	1,414214
PB	2,828427
QB	2,236068
RB	1
SB	2
TB	2,236068

Tabel 2 di atas merupakan hasil perhitungan jarak iterasi pertama untuk $x_p=3$; $y_p=2$. Dari hasil tersebut kemudian kita ikuti langkah algoritma kedua yaitu Ambil semua point yang *density reachable* terhadap titik pusatnya. Dimana *density reachable* didapat dari sekumpulan titik *item directly density-reachable* dengan syarat titik pusat dengan titik lainnya memiliki jarak tidak lebih atau sam dengan dari nilai *Eps* yang ditentukan. Karena $Eps=2$ maka nilai titik yang memenuhi syarat adalah.

Tabel 3. Titik yang memenuhi syarat iterasi I

Titik	X	Y
A	2	1
B	3	2
C	2	1
D	3	3
F	4	1
H	3	4

Titik	X	Y
I	1	2
M	3	5
O	4	6
R	2	2
S	1	2

Dari jumlah titik yang terpilih tersebut yaitu berjumlah 12. Jumlah ini sudah memenuhi untuk terbentuknya *neighborhood core object* karena jumlah objek *e- neighborhood* sudah memenuhi jumlah $MinPts=3$ dimana syarat terbentuknya *neighborhood core object* adalah sebuah objek berisi setidaknya jumlah minimal, $MinPts$. Sehingga grafik titik bisa digambarkan sebagai berikut untuk memberikan tanda dari hasil dari iterasi I Untuk menentukan *core point*(titik pusat) pada iterasi selanjutnya, dengan cara memilih titik yang memiliki jarak terjauh yang masih termasuk dalam dari *core object* pada iterasi pertama.

Tabel 4. Titik terpilih dengan masing-masing jarak iterasi I

Titik	XY	Jarak Ke titik B
A	2 1	1,414214
B	3 2	0
C	2 1	1,414214
D	3 3	1,414214
F	4 1	1,414214
H	3 4	2
I	1 2	2
M	3 5	1
O	4 6	2,236068
R	2 2	1,414214
S	1 2	2,828427

Dari tabel di atas titik H terpilih sebagai titik pusat untuk iterasi II dengan koordinat $x=3$, $y=4$ menjadi titik pusatnya.

Iterasi II

Iterasi II sama dengan iterasi sebelumnya yaitu menghitung jarak masing masing *point* atau titik terhadap titik pusat H. Berikut adalah perhitungan jarak masing-masing titik dengan *core point* untuk iterasi kedua, pada soal diatas dengan menggunakan rumus jarak diatas:

- $MinPts : 3$
- $Eps : 2$
- Pusat titik H ; $x_p=3$; $y_p=4$ Perhitungannya

$$AH = \sqrt{(2-3)^2 + (1-4)^2} = 3,162278$$

Dilakukan hal yang sama untuk titik yang lain, sehingga didapatkan perhitungan jarak untuk iterasi II sebagai berikut

Tabel 5. Hasil Jarak Iterasi II

Jarak	Hasil
AH	3,162278
BH	2
CH	3,162278
DH	1
EH	3,605551
FH	3,162278
GH	2,236068
HH	0
IH	2,828427
JH	2
KH	2,236068
LH	1
MH	2,236068
NH	3,605551
OH	3,162278
PH	2
QH	1

Jarak	Hasil
RH	2,236068
SH	2,828427
TH	3,605551

Tabel 5 diatas merupakan hasil perhitungan jarak iterasi pertama untuk $x_p=3$; $y_p=4$. Dari hasil tersebut kemudian kita ikuti langkah algoritma kedua yaitu Ambil semua point yang *density reachable* terhadap titik pusatnya. Dimana *density reachable* didapat dari sekumpulan titik *item directly density-reachable* dengan syarat titik pusat dengan titik lainnya memiliki jarak tidak lebih atau sam dengan dari nilai *Eps* yang ditentukan. Karena $Eps=2$ maka nilai titik yang memenuhi syarat adalah

Tabel 6. Titik yang memenuhi syarat iterasi II

Titik	X	Y
B	3	2
D	3	3
H	3	4
J	1	4
L	2	4
P	1	4
Q	2	4

Dari jumlah titik yang terpilih tersebut yaitu berjumlah 7. Jumlah ini sudah memenuhi untuk terbentuknya *neighborhood core object* karena jumlah objek e- *neighborhood* sudah memenuhi jumlah $MinPts=3$ dimana syarat terbentuknya *neighborhood core object* adalah sebuah objek berisi setidaknya jumlah minimal, $MinPts$. Sehingga grafik titik bisa digambarkan sebagai berikut untuk memberikan tanda dari hasil dari iterasi II. Untuk menentukan *core point*(titik pusat) pada iterasi selanjutnya, dengan cara memilih titik yang memiliki jarak terjauh yang masih termasuk dalam dari *core object* pada iterasi pertama.

Tabel 7. Titik terpilih dengan masing-masing jarak iterasi II

Titik	X	Y	Jarak Ke titik H
B	3	2	2
D	3	3	1
H	3	4	0
J	1	4	2
L	2	4	1
P	1	4	2
Q	2	4	1

Dari tabel 7 di atas titik Q terpilih sebagai titik pusat untuk iterasi III dengan koordinat $x=2$, $y=4$ menjadi titik pusatnya.

Iterasi III

Iterasi III sama dengan iterasi sebelumnya yaitu mengitung jarak masing masing *point* atau titik terhadap titik pusat Q. Berikut adalah perhitungan jarak masing-masing titik dengan *core point* untuk iterasi kedua, pada soal diatas dengan menggunakan rumus jarak diatas:

- $MinPts : 3$
- $Eps : 2$
- Pusat titik Q ; $x_p=2$; $y_p=4$ Perhitungannya

$$AQ = \sqrt{(2 - 2)^2 + (1 - 4)^2} = 3$$

Dilakukan hal yang sama untuk titik yang lain, sehingga didapatkan perhitungan jarak untuk iterasi III sebagai berikut

Tabel 8. Hasil Jarak Iterasi III

Jarak	Hasil
AQ	3
BQ	2,236068
CQ	3
DQ	1,414214
EQ	4,242641
FQ	3,605551
GQ	1,414214
HQ	1

Jarak	Hasil
IQ	2,236068
JQ	1
KQ	1,414214
LQ	0
MQ	2
NQ	3,162278
OQ	3,605551
PQ	1
QQ	0
RQ	2
SQ	2,236068
TQ	3,162278

Tabel 8 diatas merupakan hasil perhitungan jarak iterasi pertama untuk $x_p=2$; $y_p=4$. Dari hasil tersebut kemudian kita ikuti langkah algoritma kedua yaitu Ambil semua point yang *density reachable* terhadap titik pusatnya. Dimana *density reachable* didapat dari sekumpulan titik *item directly density-reachable* dengan syarat titik pusat dengan titik lainnya memiliki jarak tidak lebih atau sam dengan dari nilai *Eps* yang ditentukan. Karena $Eps=2$ maka nilai titik yang memenuhi syarat adalah.

Tabel 9. Titik yang memenuhi syarat iterasi III

Titik	X	Y
D	3	3
G	1	3
H	3	4
J	1	4
L	2	4
M	2	2
P	1	4
Q	2	4
R	2	2

Dari jumlah titik yang terpilih tersebut yaitu berjumlah 9. Jumlah ini sudah memenuhi untuk terbentuknya *neighborhood core object* karena jumlah objek *e- neighborhood* sudah memenuhi jumlah $MinPts=3$ dimana syarat terbentuknya *neighborhood core object* adalah sebuah objek berisi setidaknya jumlah minimal, $MinPts$. Sehingga grafik titik bisa digambarkan sebagai berikut untuk memberikan tanda dari hasil dari iterasi III. Untuk menentukan *core point*(titik pusat) pada iterasi selanjutnya, dengan cara memilih titik yang memiliki jarak terjauh yang masih, termasuk dalam dari *core object* pada iterasi pertama.

Tabel 10. Titik terpilih dengan masing-masing jarak iterasi III

Titik	X	Y	Jarak Ke titik Q
D	3	3	1,414214
G	1	3	1,414214
H	3	4	1
J	1	4	1
L	2	4	1,414214
M	2	2	2
P	1	4	1
Q	2	4	0
R	2	2	2

Dari tabel 10 di atas titik Q terpilih sebagai titik pusat untuk iterasi III dengan koordinat $x=3$, $y=4$ menjadi titik pusatnya.

Tabel 11. hasil iterasi I sampai iterasi III

Iterasi I	Iterasi II	Iterasi III
A	B	D
B	D	G
C	H	H
D	J	J
F	L	L
H	P	M

Iterasi I	Iterasi II	Iterasi III
I	Q	P
M		Q
O		R
R		
S		

Berdasarkan dari pengelompokkan dari iterasi I-III terdapat variabel yang muncul dua kali setiap iterasi yaitu variabel B, D, H, M, Q dan R. Dimana Variabel B adalah mahasiswa berasal dari daerah Nias, Variabel D adalah mahasiswa berasal dari daerah Deli Serdang, Variabel H adalah mahasiswa berasal dari daerah Aceh, Variabel M adalah mahasiswa berasal dari daerah Nias, Variabel Q adalah mahasiswa berasal dari daerah Aceh, Variabel R adalah mahasiswa berasal dari daerah Nias. Maka Dapat di Simpulkan bahwa dari iterasi I-III mahasiswa yang paling banyak berasal dari daerah Nias.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan menggunakan metode DBSCAN terhadap data asal wilayah mahasiswa Universitas Budi Darma Medan tahun 2018–2019, dapat disimpulkan bahwa proses clustering berhasil mengelompokkan mahasiswa berdasarkan kepadatan wilayah asalnya. Dari tabel hasil preprocessing (Tabel 1) terlihat bahwa mahasiswa berasal dari berbagai sekolah dan wilayah yang tersebar, namun setelah melalui proses iterasi clustering (Tabel 2 hingga Tabel 11), ditemukan bahwa beberapa titik data atau variabel lebih sering muncul sebagai core point maupun titik terhubung dalam setiap iterasi. Variabel-variabel yang berulang di antaranya adalah B, D, H, M, Q, dan R yang menunjukkan asal mahasiswa dari wilayah tertentu seperti Nias, Deli Serdang, dan Aceh. Fakta bahwa variabel B, M, dan R yang mewakili mahasiswa asal Nias muncul lebih dominan dalam tiga iterasi membuktikan bahwa mahasiswa dari Nias merupakan kelompok terbesar di antara asal wilayah mahasiswa lainnya. Dengan demikian, dapat disimpulkan bahwa sebaran mahasiswa Universitas Budi Darma Medan memiliki konsentrasi paling tinggi dari daerah Nias, disusul oleh wilayah lain seperti Aceh dan Deli Serdang. Hasil ini menunjukkan bahwa DBSCAN efektif digunakan untuk mengidentifikasi pola distribusi asal mahasiswa secara lebih terstruktur dan dapat menjadi dasar strategi promosi kampus yang lebih tepat sasaran, terutama dengan memperhatikan potensi besar wilayah Nias sebagai penyumbang mahasiswa terbanyak. Kesimpulan ini juga menegaskan pentingnya penerapan teknik clustering dalam analisis data pendidikan karena mampu memberikan gambaran jelas mengenai sebaran mahasiswa, sekaligus mendukung kebijakan institusi dalam mengoptimalkan perencanaan pengembangan dan promosi di masa yang akan datang.

REFERENCES

- [1] A. Jayadi, R. D. A. Saputra, A. T. Sofyan, dan S. Hidayat, “Strategi Promosi Program Studi Melalui Pendekatan Kolaboratif,” *Oikos Nomos J. Kaji. Ekon. dan Bisnis*, vol. 16, no. 2, hal. 80–91, 2023, doi: <https://doi.org/10.37479/jkeb.v16i2.23280>.
- [2] H. Du, S. Chen, H. Niu, dan Y. Li, “Application of DBSCAN clustering algorithm in evaluating students’ learning status,” in *2021 17th International Conference on Computational Intelligence and Security (CIS)*, 2021, hal. 372–376. doi: 10.1109/CIS54983.2021.00084.
- [3] V. E. Montano dan S. F. Schneider, “Unveiling the Patterns of Employability in Higher Education: A DBSCAN Cluster Analysis of College Programs,” *Eur. J. Contemp. Educ. E-Learning*, vol. 2, no. 6, hal. 73–89, 2024, doi: [https://doi.org/10.59324/ejceel.2024.2\(6\).05](https://doi.org/10.59324/ejceel.2024.2(6).05).
- [4] G. Wu, L. Cao, H. Tian, dan W. Wang, “HY-DBSCAN: A hybrid parallel DBSCAN clustering algorithm scalable on distributed-memory computers,” *J. Parallel Distrib. Comput.*, vol. 168, hal. 57–69, 2022, doi: <https://doi.org/10.1016/j.jpdc.2022.06.005>.
- [5] W. P. N. Kertanah, S. R. Aini, L. M. Amrullah, dan M. Sya’roni, “Comparison of Algorithms K-Means and DBSCAN for Clustering Student Cognitive Learning Outcomes in Physics,” *KAPPA J. Phys. Phys. Educ.*, vol. 7, no. 2, hal. 1–8, 2023, doi: <https://doi.org/10.29408/kpj.v7i2.18428>.
- [6] D. Deng, “DBSCAN clustering algorithm based on density,” in *2020 7th international forum on electrical engineering and automation (IFEEA)*, 2020, hal. 949–953. doi: 10.1109/IFEEA51475.2020.00199.
- [7] I. Sentra, I. Berbasis, dan G. Map, “Implementasi Metode Clustering dengan Algoritma DBSCAN Untuk Identifikasi Sentra Industri Berbasis Google Map,” *G-Tech J. Teknol. Terap.*, vol. 8, no. 3, hal. 2112–2121, 2024, doi: <https://doi.org/10.33379/gtech.v8i3.4959>.
- [8] E. Riswanto, “Pengelompokkan Jumlah Kunjungan Mahasiswa ke Perpustakaan Kampus Menggunakan Algoritma DBSCAN,” *G-Tech J. Teknol. Terap.*, vol. 7, no. 1, hal. 75–81, 2023, doi: <https://doi.org/10.33379/gtech.v7i1.1925>.
- [9] R. Anggara dan A. Rahman, “Implementasi Algoritma DBSCAN dalam Mengelompokkan Data Pasien Terdiagnosa Penyakit Ginjal Kronis (PGK),” *J. Algoritma*, vol. 3, no. 1, hal. 114–123, 2022, doi: <https://doi.org/10.35957/algoritme.v3i1.3593>.



- [10] S. A. A. Kharis dan A. H. A. Zili, “Learning Analytics dan Educational Data Mining pada Data Pendidikan,” *J. Ris. Pembelajaran Mat. Sekol.*, vol. 6, no. 1, hal. 12–20, 2022, doi: <https://doi.org/10.21009/jrpms.061.02>.
- [11] I. Khoeri dan D. I. Mulyana, “Implementasi machine learning dengan decision tree algoritma C4. 5 dalam penerimaan karyawan baru pada PT. Gitareksa dinamika jakarta,” *J. Sos. Teknol.*, vol. 1, no. 7, hal. 615–623, 2021, doi: <https://doi.org/10.59188/jurnalsostech.v1i7.126>.
- [12] S. D. Putri dan S. Sitohang, “Analisis Pola Pembelian Konsumen Menggunakan Algoritma Apriori,” *Comput. Sci. Ind. Eng.*, vol. 9, no. 7, hal. 889–898, 2023, doi: <https://doi.org/10.33884/comasiejournal.v9i7.7889>.
- [13] Z. Xing dan W. Zhao, “Block-diagonal guided DBSCAN clustering,” *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 11, hal. 5709–5722, 2024, doi: [10.1109/TKDE.2024.3401075](https://doi.org/10.1109/TKDE.2024.3401075).
- [14] O. Kulkarni dan A. Burhanpurwala, “A survey of advancements in DBSCAN clustering algorithms for big data,” in *2024 3rd International conference on Power Electronics and IoT Applications in Renewable Energy and its Control (PARC)*, 2024, hal. 106–111. doi: [10.1109/PARC59193.2024.10486339](https://doi.org/10.1109/PARC59193.2024.10486339).
- [15] F. Khotimah, R. S. Sofia, dan A. R. Maulana, “Analisis Penerapan Promosi Terhadap Keputusan Pembelian Mie Sakera Kalisat,” *J. Econ. Manag. Business, Entrep.*, vol. 2, no. 1, hal. 40–47, 2024, doi: <https://doi.org/10.31537/jembe.v2i1.2245>.
- [16] M. Hanif, “Strategi Humas untuk Promosi Penerimaan Peserta Didik Baru di SMP Islam Al Irsyad Gandrungmangu,” *Al-Ihda’ J. Pendidik. dan Pemikir.*, vol. 18, no. 1, hal. 926–939, 2023, doi: <https://doi.org/10.55558/alihda.v18i1.134>.