

Pengelompokkan Jumlah Kunjungan Mahasiswa ke Perpustakaan Kampus Menggunakan Algoritma DBSCAN

Andri Syafrianto^{1✉}, Eko Riswanto²

^{1,2} Program Studi Informatika, STMIK EL RAHMA Yogyakarta, Indonesia

Informasi Artikel

Riwayat Artikel

Diserahkan : 28-11-2022

Direvisi : 29-11-2022

Diterima : 01-12-2022

Kata Kunci:

Pengelompokkan,
DBSCAN, Kunjungan
Perpustakaan, Mahasiswa

Keywords :

*Clustering, DBSCAN,
Library Visits, Students.*

Corresponding Author:

Eko Riswanto

Program Studi Teknik Informatika, STMIK EL RAHMA Yogyakarta

Jl Karangajen No.76 Brontokusuman, Yogyakarta

Email: riswantoeko@stmikelrahma.ac.id

ABSTRAK

Perpustakaan STMIK EL Rahma Yogyakarta merupakan salah satu fasilitas penunjang akademik bagi mahasiswa untuk mendukung kegiatan KBM yang dilakukan. Adanya korelasi yang positif antara jumlah kunjungan mahasiswa ke perpustakaan terhadap peningkatan kinerja mahasiswa yang menjadikan perpustakaan sebagai sumber belajar menjadi salah satu alasan peneliti melakukan penelitian ini. Berdasarkan hal tersebut pada penelitian ini akan mencoba mengelompokkan mahasiswa STMIK EL Rahma Yogyakarta berdasarkan IPK dan jumlah kunjungan ke perpustakaan dengan algoritma DBSCAN. Adapun manfaat dari penelitian ini yaitu memberikan gambaran mengenai minat baca mahasiswa terhadap perpustakaan kampus. Berdasarkan hasil analisa terhadap 75 mahasiswa dengan nilai epsilon 0,4 dan MinPts 2 maka metode DBSCAN mampu membentuk cluster sebanyak 7 dengan 5 data noise.

ABSTRACT

The STMIK EL Rahma Yogyakarta library is one of the academic support facilities for students to support their teaching and learning activities. The existence of a positive correlation between the number of student visits to the library and the increase in student performance which makes the library a source of learning is one of the reasons researchers conducted this research. Based on this, this study will try to group STMIK EL Rahma Yogyakarta students based on GPA and number of visits to the library with the DBSCAN algorithm. The benefit of this research is to provide an overview of students' reading interest in the campus library. Based on the results of the analysis of 75 students with epsilon values of 0.4 and MinPts 2, the DBSCAN method is able to form 7 clusters with 5 noise data.

PENDAHULUAN

Pada tahun 2019, *Program for International Student Assessment (PISA)* yang di rilis *Organization for Economic Co-operation and Development (OECD)* menempatkan Indonesia di peringkat ke-62 dari 70 negara berdasarkan tingkat literasi masyarakatnya[1]. Angka ini menunjukkan bahwa Indonesia bukanlah negara yang masyarakatnya memiliki minat dan kegemaran membaca. Kebiasaan masyarakat Indonesia lebih senang menonton dan melihat

dibandingkan membaca buku adalah salah satu faktor yang menyebabkan Indonesia berada di peringkat 8 terbawah.

Di lingkungan perguruan tinggi, perpustakaan memiliki peran strategis dalam sistem pendidikan, karena perpustakaan merupakan salah satu unsur penunjang bagi mahasiswa untuk mencari ilmu pengetahuan. Keberhasilan siswa dalam pendidikan dapat diukur dengan Indeks Prestasi Kumulatif (IPK). Dimana tinggi rendahnya IPK tersebut dapat dipengaruhi oleh faktor aktivitas akademik, faktor fasilitas pendukung, dan faktor lingkungan. Faktor-faktor dalam kegiatan akademik meliputi alokasi waktu belajar, waktu yang dihabiskan untuk kegiatan ekstrakurikuler, jumlah buku pegangan yang dimiliki, jumlah mata kuliah yang diambil, dan jumlah semester aktif yang dilalui. Faktor pendukung secara ilmiah adalah frekuensi kunjungan ke perpustakaan. Faktor lingkungan tempat tinggal meliputi jarak dari rumah ke kampus dan waktu tempuh dari rumah ke kampus.

Beberapa penelitian terdahulu menyatakan bahwa adanya korelasi positif yang cukup tinggi antara frekuensi kunjungan mahasiswa ke perpustakaan dengan nilai Indeks Prestasi Kumulatif (IPK) yang didapatkan melalui pendekatan korelasi Pearson (Nisa & Aisyah, 2021). Selain itu mahasiswa yang menggunakan perpustakaan sebagai sumber belajar dapat memberikan dampak yang signifikan terhadap peningkatan kinerja mahasiswa (Liana Saputri, 2019). Merujuk pada kedua penelitian tersebut, maka artikel ini akan mencoba untuk mengelompokkan mahasiswa STMIK El-Rahma berdasarkan IPK dan frekuensi kunjungan ke perpustakaan dengan algoritma DBSCAN. Hasil dari pengelompokan dapat menjadi rujukan bagi pihak Perguruan Tinggi untuk mendapatkan gambaran mengenai kondisi dan minat baca mahasiswa terhadap perpustakaan kampus.

Penelitian mengenai DBSCAN pernah digunakan untuk mengetahui minat dalam penggunaan transportasi umum (Kristianto, 2021), mengelompokkan pola dari kasus covid-19 di dunia (Adha et al., 2021), serta mengelompokkan mahasiswa yang berhak menerima bantuan Yayasan (Indini et al., 2022). Pada penelitian ini peneliti memilih algoritma DBSCAN karena algoritma ini mampu mendapatkan cluster yang berubah-ubah dan memiliki bentuk tidak tentu secara efektif (Harjanto et al., 2021). Selain itu juga algoritma DBSCAN akan menentukan clusternya dengan kepadatan yang tinggi (Hu et al., 2021).

METODE PENELITIAN

Data yang digunakan dalam penelitian ini diperoleh dengan mengumpulkan 75 data mahasiswa S1 Informatika semester III STMIK EL-RAHMA Yogyakarta yang melakukan kunjungan ke Perpustakaan STMIK EL-RAHMA Yogyakarta selama seminggu. Algoritma DBSCAN (*Density Based Clustering*) mampu mengelompokkan data berbasis kepadatan serta untuk membedakan *noise* dari database spasial yang besar (Cinderatama et al., 2022). DBSCAN menerima dua parameter yaitu Eps (*radius*) dan MinPts (titik minimum-ambang batas) (Zhu et al., 2021). DBSCAN didasarkan pada pendekatan berbasis pusat (Supriyadi et al., 2021), di mana kepadatan diperkirakan dengan menghitung jumlah titik dalam radius tertentu di dataset yang dikenal dengan Eps. Algoritma DBSCAN memungkinkan untuk mengklasifikasikan suatu titik sebagai titik inti, titik batas dan titik kebisingan. Ide utamanya adalah bahwa untuk setiap titik dari sebuah cluster, lingkungan dengan radius tertentu (Eps) harus mengandung setidaknya sejumlah titik minimum (MinPts).

Langkah-langkah dalam metode Algoritma DBSCAN adalah (Hu et al., 2021):

1. Pilih poin secara *random* (p)
2. Ambil semua titik yang dapat dijangkau kepadatannya dari p dan sesuai dengan nilai epsilon (Eps) dan *minimum points* (MinPts). Nilai epsilon digunakan untuk menentukan jarak radius antara titik-titik dalam dataset dengan lingkungan sekitarnya. Sedangkan minimum points digunakan untuk mendapatkan jumlah minimum titik yang ada disekitar radius epsilon (Pakuani & Kurniawan, 2021).
3. Jarak antar titik dihitung dengan menggunakan euclidean distance

4. Sebuah cluster terbentuk, dengan p adalah titik intinya
5. Kunjungi titik berikutnya dari kumpulan data, jika p adalah titik batas dan tidak ada titik yang kepadatannya dapat dicapai dari p
6. Ulangi proses di atas sampai semua poin telah diperiksa

HASIL DAN PEMBAHASAN

Jumlah data yang digunakan dalam dataset berjumlah 75 mahasiswa. Data asli mahasiswa, IPK dan jumlah kunjungan ke perpustakaan terdapat pada tabel 1.

Tabel 1. Tabel dataset

Mahasiswa ke-	IPK	Jumlah Kunjungan
1	3	2
2	3,83	3
3	3,75	3
4	3	1
5	1,92	1
6	3,6	2
7	1,5	1
8	2,98	1
9	3,33	2
10	3,75	1
...
75	3,43	2

Dari dataset tersebut akan diambil mahasiswa ke-10 sebagai *centroid* pada iterasi pertama dengan IPK 3,75 dan jumlah kunjungan sebanyak 1 kali. Selanjutnya DBSCAN akan menghitung jarak antara mahasiswa nomor 1, 2, 3, dan 4 dengan menggunakan *Euclidean Distance*. Rangkuman mengenai hasil perhitungan ini terdapat pada Tabel 2.

Tabel 2. Tabel hasil perhitungan jarak dengan titik pusat

Mahasiswa 10 ke-i	Jarak
1	1,75
2	4,08
3	4
4	0,75
5	1,83
6	1,15
7	2,25
8	0,77
9	1,42
10	0
...	...
75	1,32

Selanjutnya akan mengambil semua titik yang dapat dijangkau kepadatannya dari titik pusat (p) dan menetapkan epsilon (Eps) serta minimum points (MinPts). Pada penelitian ini, nilai epsilon ditetapkan pada skor 10 dan minimum points peneliti tetapkan pada skor 2. Berdasarkan nilai kedua parameter tersebut, *density reachable* titik yang berada dalam jangkauan nilai epsilon = 10 ditunjukkan pada tabel 3.

Tabel 3. Tabel *density reachable* titik pusat dengan epsilon

Mahasiswa ke-i dan 10	Jarak
10.10	0
71.10	0,5
70.10	0,54
63.10	0,58
72.10	0,58
73.10	0,58
4.10	0,75
12.10	0,75
62.10	0,75
8.10	0,77

Berdasarkan tabel 3, semua nilai kurang dari nilai parameter MinPts. Oleh karenanya nilai pada tabel 3 memenuhi sebagai *Neighborhood Core Object*. Selanjutnya pada iterasi yang kedua, nilai titik pusat diambil berdasarkan titik yang memiliki jarak terjauh dengan titik pusat di iterasi yang sebelumnya. Berdasarkan tabel 3, nilai titik pusat pada iterasi yang kedua adalah mahasiswa ke-8. Iterasi ini akan berulang sampai dengan iterasi ke-5 dengan nilai *density reachable* titik pusatnya adalah mahasiswa ke-69 dengan nilai epsilon 10. Rangkuman mengenai nilai density ini terdapat pada table 4.

Tabel 4. Tabel *density reachable* titik pusat mahasiswa ke-69 dengan epsilon 10

Mahasiswa ke-i dan 10	Jarak
69	0
56	0,01
8	0,06
4	0,08
12	0,08
62	0,08
35	0,17
63	0,25
72	0,25
73	0,25

Berdasarkan table 4 tersebut, maka semua nilai berada kurang dari nilai parameter MinPts. Oleh karenanya nilai pada tabel 4 memenuhi sebagai *Neighborhood Core Object*. Nilai Density reachable pada iterasi yang kelima sudah konvergen karena titik terjauh (mahasiswa ke-73) sama dengan titik terjauh pada iterasi yang keempat. Karena sudah konvergen, maka proses perhitungan dihentikan. Hasil titik pada iterasi yang pertama hingga iterasi kelima terangkum pada tabel 5.

Tabel 5. Hasil titik untuk 5 iterasi dengan parameter epsilon=10 dan MinPts=2

Mahasiswa ke-i dan 10	Jarak
69	0
56	0,01
8	0,06
4	0,08
12	0,08
62	0,08
35	0,17
63	0,25
72	0,25
73	0,25

Berdasarkan table 5 diketahui bahwa semua nilai sudah kurang dari nilai parameter MinPts yang ditentukan. Oleh karena itu nilai pada tabel 5 sudah memenuhi nilai *Neighborhood*

Core Object. Selain itu nilai *Density Reachable* pada iterasi yang kelima juga sudah konvergen karena titik terjauh (mahasiswa ke-73) sama dengan titik terjauh pada iterasi keempat. Karena sudah konvergen, maka proses perhitungan dihentikan. Adapun hasil titik pada iterasi yang pertama hingga iterasi yang kelima disajikan dalam tabel 6.

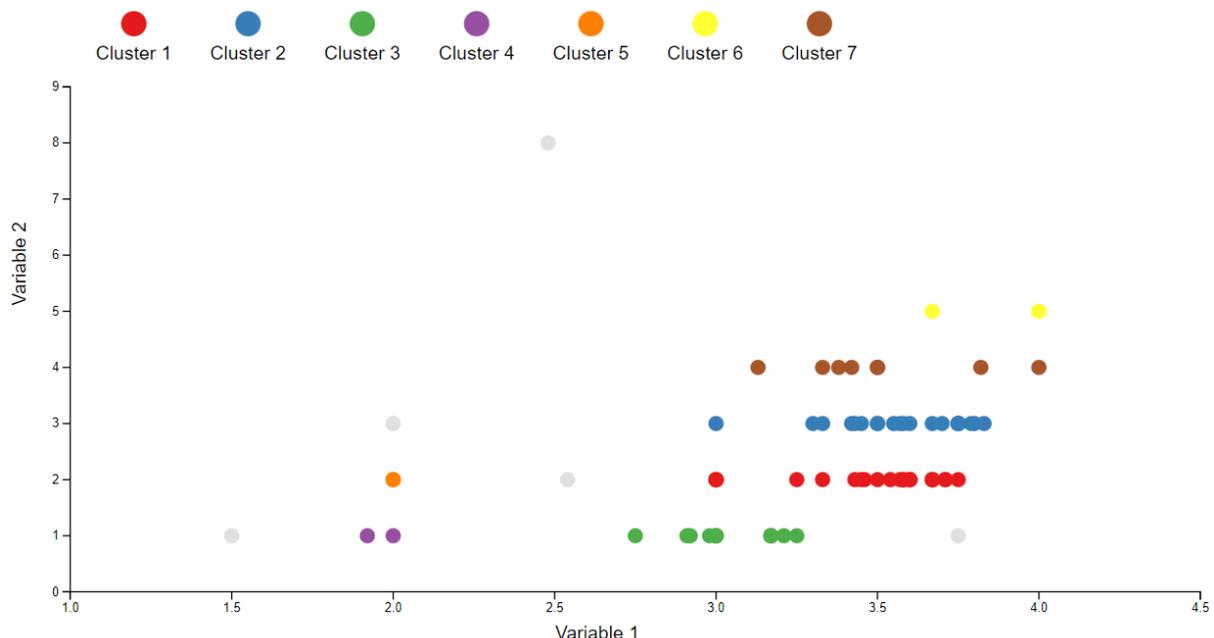
Tabel 6. Hasil titik untuk 5 iterasi dengan parameter $\epsilon=10$ dan $\text{MinPts}=2$

10	8	35	63	69
71	4	56	72	56
70	12	69	73	8
63	62	8	70	4
72	69	4	71	12
73	56	12	4	62
4	63	62	12	35
12	72	63	62	63
62	73	72	8	72
8	35	73	69	73

Berdasarkan hasil analisis dan Tabel 6 didapatkan bahwa dengan menggunakan parameter $\epsilon=10$ dan $\text{MinPts}=2$ dataset belum dapat mengelompok dengan baik, hal ini disebabkan karena hampir 90% titik belum terpisah secara jelas kelompoknya. Oleh karena itu perlu dilakukan perubahan terhadap nilai ϵ dan MinPts .

Pengujian

Hasil uji coba pada parameter ϵ dan MinPts menggunakan bahasa pemrograman R, didapatkan nilai stabil pada $\epsilon=0.4$ dan $\text{MinPts}=2$. Pengelompokan grafik pada dataset yang diajukan dapat dilihat pada gambar 1.



Gambar 1. Hasil Pengelompokkan

Berdasarkan Gambar 1, maka Variabel 1 adalah jumlah kunjungan dan variabel 2 adalah IPK. Dari hasil percobaan yang ditampilkan di gambar 1, didapatkan 7 cluster dan data noise sebanyak 4 data. Berdasarkan hal tersebut dapat disimpulkan bahwa metode DBSCAN cenderung

melakukan pengelompokan berdasarkan jumlah kunjungan. Mahasiswa yang berkunjung ke perpustakaan 2 dan 4 kali memiliki IPK yang lebih baik dibandingkan kelompok lainnya.

Sedangkan untuk mengukur performa DBSCAN dalam melakukan pengelompokan data kunjungan mahasiswa ke perpustakaan STMIK EL RAHMA terdapat pada Tabel 7.

Table 7. hasil pengujian DBSCAN

Cluster ke-	Nilai DBI
1	0,98
2	0,86
3	0,92
4	0,88
5	0,82
6	0,80
7	0,79

Berdasarkan Tabel 7 maka didapatkan bahwa nilai DBI untuk algoritma DBSCAN terdapat pada cluster ke-7 dengan nilai DBI sebesar 0,79

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian diatas, maka dapat disimpulkan bahwa Metode DBSCAN dapat dengan baik melakukan pengelompokan mahasiswa berdasarkan 2 parameter yaitu jumlah kunjungan ke perpustakaan dan IPK. Selain itu dengan memberikan dua nilai yaitu $\epsilon=0.4$ dan $MinPts=2$, maka metode DBSCAN membentuk cluster sebanyak 7 pada dataset dengan 5 data noise dari ujicoba yang dilakukan. Selain itu dari hasil Analisa yang dilakukan DBSCAN cenderung melakukan pengelompokan berdasarkan jumlah kunjungan mahasiswa ke perpustakaan. Dengan mahasiswa yang berkunjung ke perpustakaan 2 dan 4 kali memiliki IPK yang lebih baik dibandingkan kelompok lainnya

Saran

Pengelompokan mahasiswa berdasarkan 2 parameter yaitu IPK dan jumlah kunjungan membuat metode DBSCAN kurang bervariasi dalam penentuan jumlah cluster. Ke depannya jumlah waktu yang dihabiskan mahasiswa selama di perpustakaan mungkin bisa ditambahkan guna mendapatkan cluster yang lebih bervariasi.

REFERENSI

- Adha, R., Nurhaliza, N., Sholeha, U., & Mustakim, M. (2021). Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia. *SITEKIN: Jurnal Sains, Teknologi Dan Industri*, 18(2), 206–211.
- Cinderatama, T. A., Zulmy Alhamri, R., Yunhasnawa, Y., Polinema, P., Kediri, K., Maskumambang, J. L., Mojoroto, K., Informasi, J. T., Malang, N., Soekarno Hatta, J., Lowokwaru, K., & Malang, K. (2022). Implementasi Metode K-Means, Dbscan, dan Meanshift Untuk Analisis Jenis Ancaman Jaringan Pada Intrusion Detection System. *Jurnal Inovtek Polbeng - Seri Informatika*, 7(1), 169–184. <https://doi.org/10.35314/isi.v7i1.2336>
- Harjanto, T. D., Vatesia, A., & Faurina, R. (2021). Analisis Penetapan Skala Prioritas Penanganan Balita Stunting Menggunakan Metode DBSCAN Clustering (Studi Kasus Data Dinas Kesehatan Kabupaten Lebong). *Rekursif: Jurnal Informatika*, 9(1), 30–42. <https://doi.org/10.33369/rekursif.v9i1.14982>
- Hu, L., Liu, H., Zhang, J., & Liu, A. (2021). KR-DBSCAN: A density-based clustering algorithm based on reverse nearest neighbor and influence space. *Expert Systems with Applications*,

186(August), 115763. <https://doi.org/10.1016/j.eswa.2021.115763>

- Indini, D. P., Siburian, S. R., & Utomo, D. P. (2022). Implementasi Algoritma DBSCAN untuk Clustering Seleksi Penentuan Mahasiswa yang Berhak Menerima Beasiswa Yayasan. *SENASHTEK*, 325–331.
- Kristianto, A. (2021). Analisa Performa K-Means dan DBSCAN dalam Clustering Minat Penggunaan Transportasi Umum. *Elkom : Jurnal Elektronika Dan Komputer*, 14(2), 368–372. <https://doi.org/10.51903/elkom.v14i2.551>
- Liana Saputri, R. (2019). Pengaruh Pemanfaatan Perpustakaan Sebagai Sumber Belajar Terhadap Prestasi Belajar Mahasiswa Pendidikan Guru Madrasah Ibtidaiyah (PGMI) Semester VII Fakultas Tarbiyah dan Keguruan Tahun Akademik 2019/2020 [Universitas Islam Negeri Mataram]. In *Universitas Islam Negeri Mataram*. <http://etheses.uinmataram.ac.id/1833/>
- Nisa, M. H., & Aisyah, A. F. (2021). Korelasi Pemanfaatan Koleksi Digital Perpustakaan dengan Prestasi Mahasiswa di Universitas Muhammadiyah Surakarta. <http://eprints.ums.ac.id/id/eprint/93411>
- Pakuani, K. W., & Kurniawan, R. (2021). Kajian Penentuan Nilai Epsilon Optimal Pada Algoritma DMDBSCAN Dan Pemetaan Daerah Rawan Gempa Bumi Di Indonesia Tahun 2014-2020. *Seminar Nasional Official Statistics 2021*, 2021(01), 991–1000. <https://doi.org/10.34123/semnasoffstat.v2021i1.847>
- Supriyadi, A., Triayudi, A., & Sholihati, I. D. (2021). Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas. *JIPPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 6(2), 229–240. <https://doi.org/10.29100/jipi.v6i2.2008>
- Zhu, Q., Tang, X., & Elahi, A. (2021). Application of the novel harmony search optimization algorithm for DBSCAN clustering. *Expert Systems with Applications*, 178(April), 115054. <https://doi.org/10.1016/j.eswa.2021.115054>