

# Klasterisasi Perguruan Tinggi Swasta Berdasarkan Minat Siswa Menggunakan Metode K-Medoids

Andika Dwi Cipta <sup>#1</sup>, Asep Id Hadiana <sup>\*2</sup>, Fajri Rahmat Umbara <sup>#3</sup>

Program Studi Informatika  
Universitas Jenderal Achmad Yani  
Jl. Terusan Jend. Sudirman, Cimahi, Jawa Barat 40531

<sup>1</sup> andika.dwi.@student.unjani.ac.id

<sup>2</sup> asep.hadiana@lecture.unjani.ac.id

<sup>3</sup> fajri.rakhmat@lecture.unjani.ac.id

## Abstrak

Perguruan tinggi merupakan lembaga tertinggi dari sistem pendidikan nasional. Perguruan tinggi juga memiliki peran penting dalam mengembangkan kemampuan manusia untuk dilatih dan berkembang sehingga menjadi individu yang mempunyai pola pikir dan analisis kritis dalam memecahkan sebuah masalah. Terdapat berbagai jenis perguruan tinggi di Indonesia yaitu PTN (Perguruan Tinggi Negeri), PTS (Perguruan Tinggi Swasta), PTA (Perguruan Tinggi Akademik), dan PTK (Perguruan Tinggi Kedinasan). Bagi siswa/pelajar yang akan melanjutkan pendidikan perguruan tinggi mengalami kesulitan mencari informasi perguruan tinggi mana yang diminati karena jumlah perguruan tinggi di Indonesia terhitung sangat banyak khususnya pada perguruan tinggi swasta. Kumpulan-kumpulan data umum mengenai perguruan tinggi swasta tersebut terhitung sangat banyak dan dibutuhkan suatu kesimpulan data dalam penentuan setiap Perguruan Tinggi Swasta yang diminati. Oleh sebab itu dibutuhkan pengelompokan data Perguruan Tinggi Swasta untuk dapat mengetahui informasi perguruan tinggi yang diminati. Dalam penelitian ini algoritma *K-Medoids* digunakan untuk melakukan klasterisasi. Penelitian ini menghasilkan 4 buah *cluster* dimana *cluster 2* merupakan perguruan tinggi swasta yang sangat diminati, *cluster 1* merupakan perguruan tinggi swasta yang cukup diminati, *cluster 0* merupakan perguruan tinggi swasta yang kurang diminati, dan *cluster 3* merupakan perguruan tinggi swasta yang tidak diminati. Hasil nilai evaluasi dengan menggunakan metode *Silhouette Coefficient* untuk 4 *cluster* yaitu 0,72589 sehingga hasil *cluster* yang terbentuk masuk kedalam *cluster strong structure* atau merupakan *cluster* terbaik.

**Keywords:** Klasterisasi, K-Medoids, Minat Siswa

## I. PENDAHULUAN

**P**endidikan merupakan bidang yang penting bagi kebutuhan setiap orang untuk mencapai pola hidup yang lebih baik. Pendidikan mengajarkan perilaku yang baik, mencerdaskan pola pikir dalam diri, mengembangkan kreativitas, memiliki pribadi yang bertanggung jawab dan memperluas pengetahuan. Dalam proses pendidikan, setiap orang memaksimalkan potensinya dan mengembangkannya hingga kualitas pribadinya matang agar mampu memenuhi kebutuhan lingkungan. Oleh karena itu, selain pengembangan diri manusia, pendidikan juga menentukan pola pikir dalam bermasyarakat. Maka standar kualitas setiap orang dapat diukur dari tingkat pendidikan. Kebutuhan akan kualitas dicapai melalui lembaga formal dari Pendidikan anak usia dini, sekolah

dasar, jenjang pendidikan menengah, dan jenjang pendidikan tinggi sebagai lembaga tertinggi dari sistem pendidikan nasional.

Perguruan tinggi juga memiliki peran penting dalam mengembangkan kemampuan manusia untuk dilatih dan berkembang sehingga menjadi individu yang memiliki wawasan yang baik[1]. Terdapat berbagai jenis perguruan tinggi di Indonesia yaitu PTN(Perguruan Tinggi Negeri), PTS(Perguruan Tinggi Swasta), PTA(Perguruan Tinggi Akademik), dan PTK(Perguruan Tinggi Kedinasan).

Informasi mengenai perguruan tinggi menjadi salah satu faktor terpenting dalam pemilihan perguruan tinggi, dimana pelajar yang akan melanjutkan studinya ke jenjang yang lebih tinggi yaitu perguruan tinggi umumnya mengalami kesulitan dalam memperoleh informasi-informasi perguruan tinggi khususnya informasi Perguruan Tinggi Swasta[2]. Perguruan Tinggi Swasta berlomba-lomba dalam menawarkan jasa pendidikannya dengan meningkatkan promosi serta kualitas yang dapat menarik minat calon mahasiswa dalam memilih studi di Perguruan Tinggi Swasta sesuai dengan keinginan calon mahasiswa. Faktor-faktor yang perlu diperhatikan oleh Perguruan Tinggi Swasta dalam menarik minat calon mahasiswa adalah dari sisi faktor sistem pembayaran, lingkungan Perguruan Tinggi, faktor ekonomi, fasilitas dan promosi yang lebih kreatif. Salah satu karakteristik PTS yang relatif diminati seperti laboratorium, biaya pendidikan, serta gedung perkuliahan yang permanen dan lain sebagainya[3]. Jika dilihat dari data Badan Pusat Statistik (BPS) pada tahun 2021 terdapat 2.990 unit Perguruan Tinggi Swasta yang ada di Indonesia. Oleh sebab itu dibutuhkan pengelompokan data Perguruan Tinggi, dalam data mining terdapat teknik pengelompokan atau *clustering* yang dapat menghasilkan informasi-informasi yang belum ada atau belum diketahui sebelumnya.

Data yang didapatkan dari *website* ayokuliah.id yang digunakan sebagai acuan pengelompokan data Perguruan Tinggi Swasta. Metode yang dipakai adalah Algoritma *K-Medoids* karena melihat dalam penelitian[4] kasus yang diteliti memiliki kesamaan dalam bidang pendidikan yaitu dengan studi kasus penentuan jurusan siswa pada SMA algoritma *K-Medoids* dapat mengelompokan data dalam menentukan jurusan siswa dengan baik. Sehingga algoritma *K-Medoids* dirasa cocok digunakan dalam penelitian ini. Selain itu *K-Medoids* juga dapat mengurangi noise dan outlier pada data[5].

## II. LANDASAN TEORI

### A. *Clustering*

*Clustering* adalah teknik yang banyak digunakan, yang tujuannya adalah untuk memberikan wawasan ke dalam data dengan mempartisi data (objek) menjadi kelompok-kelompok (*cluster*) yang terpisah-pisah dan homogen, sehingga objek-objek dalam sebuah pengelompokan data lebih mirip satu sama lain daripada objek-objek di dalam *cluster* lainnya[6]. *Clustering* dapat dimanfaatkan untuk mengidentifikasi ringkasan dokumen teks yang koheren, yang dapat memberikan petunjuk tentang keseluruhan isi korpus[7]. Algoritma pengelompokan telah sering dipelajari di berbagai bidang termasuk pembelajaran mesin, jaringan saraf dan statistik, itu k-berarti algoritma, pertama kali diusulkan oleh MacQueen, adalah metode pengelompokan partisi paling populer yang telah menarik minat besar dalam literatur. Tujuan dari k- algoritma adalah untuk mempartisi objek menjadi k *cluster* sehingga kesamaan dalam kelompok dimaksimalkan[8].

### B. Metode *Elbow*

Metode *Elbow* merupakan metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan cara melihat persentase perbandingan antara jumlah *cluster* yang akan membentuk siku pada sebuah titik. Metode ini memberikan ide/gagasan dengan cara memilih nilai *cluster* kemudian menjumlahkan nilai *cluster* untuk dijadikan model data dalam menentukan *cluster* terbaik. Dan selain itu persentase perhitungan yang dihasilkan merupakan perbandingan antara jumlah *cluster* yang ditambahkan[9].

$$SSE = \sum_{K=1}^K \sum_{x1 \in Ck} \|xi - ck\|_2^2 \quad (1)$$

Keterangan :

SSE : *Sum Of Square Error*

Xi : nilai atribut dari data ke 1

$C_k$  : nilai atribut titik pusat *cluster* ke 1

### C. Algoritma *K-Medoids*

Algoritma *K-Medoids* menutupi kekurangan dari *K-Means* karena *K-means* sensitif terhadap *outlier* dan penggunaan fungsi *squared-error*. Motivasi utama dalam algoritma berbasis *K-medoids* adalah bahwa mereka memilih objek aktual untuk mewakili *cluster*, menggunakan satu objek representatif per *cluster* dan menetapkan masing-masing objek yang tersisa ke objek representatif yang paling mirip. Karena *cluster* diwakili melalui perwakilan dan bukan pusat *cluster*, algoritme berbasis *K-medoids* dicirikan dengan kurang sensitif terhadap *outlier* dan nilai *ekstrem*[10]. Salah satu algoritma berbasis *K-medoids* yang paling penting adalah algoritma PAM (*Partitioning Around Medoids*). Algoritma mengambil  $k$ , jumlah *cluster* dalam kumpulan data dan kumpulan data dengan  $n$  objek sebagai parameter input dan membangun satu set  $k$  *cluster*[11].

Berikut Langkah-langkah untuk algoritma *K-medoids*:

1. Tentukan *medoid* awal dengan memilih  $K$  sebagai objek pada sekumpulan data objek sesuai jumlah titik pusat pada *cluster*.
2. Setiap objek yang dipilih atau yang bukan *medoid* di masukan ke dalam *cluster* berdasarkan jarak yang paling dekat dengan menggunakan persamaan *Euclidean Distance*.
3. Pilih objek masing-masing *cluster* secara acak untuk *medoid* baru.
4. Hitung jarak masing-masing *cluster* dengan *medoid* baru.
5. Hitung total ( $S$ ) dengan cara menghitung nilai total dari jarak baru-total jarak lama.
6. Jika dihasilkan  $S < 0$ , maka objek ditukar dengan data *cluster* untuk membentuk sekumpulan  $K$  dari objek baru sebagai *medoid*.
7. Ulangi tahap ke 3-5 sampai tidak terjadi perubahan pada *medoid*, sehingga didapatkan *cluster* objek beserta anggota *cluster* masing-masing[12].

Persamaan jarak yang digunakan untuk menghitung objek, digunakan rumus jarak *Euclidean Distance* pada persamaan berikut ini

$$d_{Euclidean(x,y)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

### D. *Silhouette Coefficient*

Evaluasi *Silhouette Coefficient* ini menggabungkan 2 metode, yaitu metode *Cohesion* dan metode *Separation*. Metode *Cohesion* digunakan untuk mengukur jarak antara objek satu dengan objek lainnya didalam sebuah *cluster*, sedangkan metode *Separation* digunakan untuk mengukur jarak dari *cluster* pertama hingga *cluster* lainnya[13]. Ada 3 tahapan dalam perhitungan *Silhouette Coefficient* yaitu diantaranya[14]:

1. Hitung rata-rata pada setiap objek  $i$  dengan semua objek yang berada dalam satu lingkup *cluster*. Sehingga akan memperoleh nilai rata-rata yang disebut  $a_i$
2. Setiap objek  $i$  nilai minimal jarak rata-rata dihitung dari titik satu ke titik lain yang berbeda *cluster*. Sehingga akan diperoleh nilai rata-rata minimum yang disebut dengan  $b_i$ .

Setelah semua nilai diketahui maka nilai *Silhouette Coefficient* dapat ditentukan dengan menggunakan persamaan berikut:

$$S_i = \frac{b_i - a_i}{\max(a_i - b_i)} \quad (3)$$

Keterangan :

$b$  : Jarak rata-rata *medoid* dengan obyek di luar *cluster*

$a$  : Jarak rata-rata antara *medoid* dengan objek di dalam

Adapun nilai *Silhouette Coefficient* pada angka antara -1 sampai dengan 1. Berikut merupakan *Silhouette Coefficient* menurut Kaufman dan Rousseeuw:

$0.7 < SC \leq 1$  tergolong kedalam *Cluster Strong Structure*

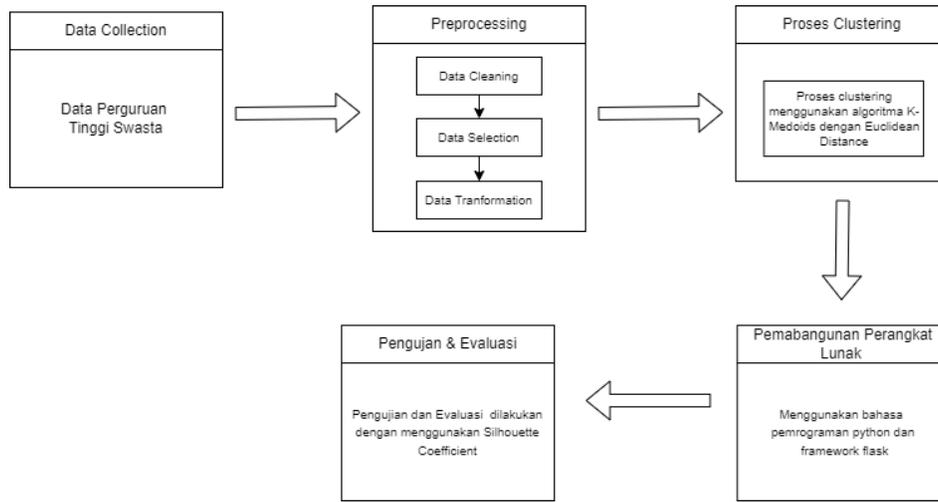
$0.5 < SC \leq 0.7$  tergolong kedalam *Cluster Medium Structure*

$0.25 < SC \leq 0.5$  tergolong kedalam *Cluster Weak Structure*

$SC \leq 0.25$  tergolong kedalam *Cluster No Structure*

## III. METODELOGI PENELITIAN

Pada proses penelitian ini dilakukan tahapan-tahapan untuk menemukan hasil yang di inginkan. Tahap penelitian digambarkan sebagai berikut.



Gambar 1 Metode Penelitian

A. Data Collection

Pada penelitian yang dilakukan data yang diambil dari *website* ayokuliah.id yaitu data Perguruan Tinggi Swasta di Pulau Jawa. Proses pengumpulan data dilakukan dengan cara manual.

B. Pre-Processing

Proses *pre-processing* dilakukan terhadap data penelitian, proses ini dilakukan agar memperbaiki data penelitian yang tidak lengkap, *noise* pada data atau data yang tidak konsisten[15].

C. Clustering

Setelah mendapatkan data penelitian, selanjutnya dilakukan proses *clustering* dengan menggunakan algoritma *K-medoids*.

D. Pembangunan Perangkat Lunak

Didalam tahapan ini merupakan tahapan implementasi penelitian yang dilakukan kedalam sebuah perangkat lunak. Dengan melakukan perancangan dan pembangunan perangkat lunak untuk mempermudah melakukan klasterisasi dan visualisasi data yang digunakan. Sistem yang akan dibangun adalah menggunakan bahasa pemrograman *python* dan *framework flask MVC*.

E. Pengujian & Evaluasi

Pada tahap ini merupakan tahap mengidentifikasi hasil *cluster* yang akan terbentuk dengan mengevaluasi nilai *Silhouette Score* menggunakan pengujian *Silhouette Coefficient*.

IV. PENGUJIAN DAN ANALISIS

Pada penelitian ini, data yang akan digunakan adalah data perguruan tinggi swasta dengan jumlah *record* 1125 data dan 35 atribut, berupa data Jumlah Mahasiswa, Jumlah Dosen, Jumlah Program Studi, Fasilitas Olahraga, Wisma/Asrama/Hotel, Teknologi Informasi, Public Area, PERS Mahasiswa, Asuransi, Gedung Pertemuan, Laboratorium, Poliklinik, Bus Kampus, Kalender Pendidikan, Pusat Pelatihan Bahasa, Perpustakaan, Sarana Ibadah, Free Hotspot, Jumlah Fasilitas, Bus Umum, Restaurant, Tempat Ibadah, Kos/Asrama, Terminal Bus, Café, Tempat Olahraga, Bandara, Mall, Rumah Sakit, Stasiun Kereta, Supermarket, Apotek, Jumlah Tempat Umum Terdekat, Akreditasi, dan Ranking Nasional.

Table 1 Dataset Perguruan Tinggi Swasta

Jumlah Mahasiswa	Jumlah Dosen	Jumlah Program Studi	Fasilitas Olahraga	Wisma	Teknologi Informasi	Public Area	...	Akreditasi
29332	538	49	Ya	Ya	Ya	Ya	...	A

17208	440	41	Ya	Ya	Ya	Ya	...	A
25754	468	34	Ya	Ya	Ya	Ya	...	A
11106	305	16	Ya	Ya	Ya	Ya	...	A
10728	332	40	Ya	Ya	Ya	Ya	...	A
...	...	...	...	...	...	...	...	...
1123	300	13	Ya	Ya	Ya	Ya	...	A

A. Pre-processing

Terdapat 3 teknik dalam preprocessing data yaitu Data *Cleaning*, Data *Transformation*, dan Data *Normalization*. Proses data *cleaning* yaitu untuk melihat ada tidaknya *missing value* pada setiap atribut. Karena data tidak ada nilai *missing value* pada setiap atribut, sehingga tidak diperlukan proses *handling missing value*.

Setelah melakukan *cleaning* dan selection selanjutnya masuk pada tahap data tranformasi yang bertujuan untuk menyesuaikan data yang akan diolah agar sesuai dengan kebutuhan. Tahapan ini mengubah variabel data dari bentuk kategorik menjadi bentuk numerik. Setelah data di transformasikan selanjutnya dilakukan proses normalisasi pada data agar data yang siap untuk proses data mining lebih baik karena data yang di normalisasi akan berada pada range 0 sampai 1 dengan menggunakan metode *scaling* yaitu *min-max normalization*.

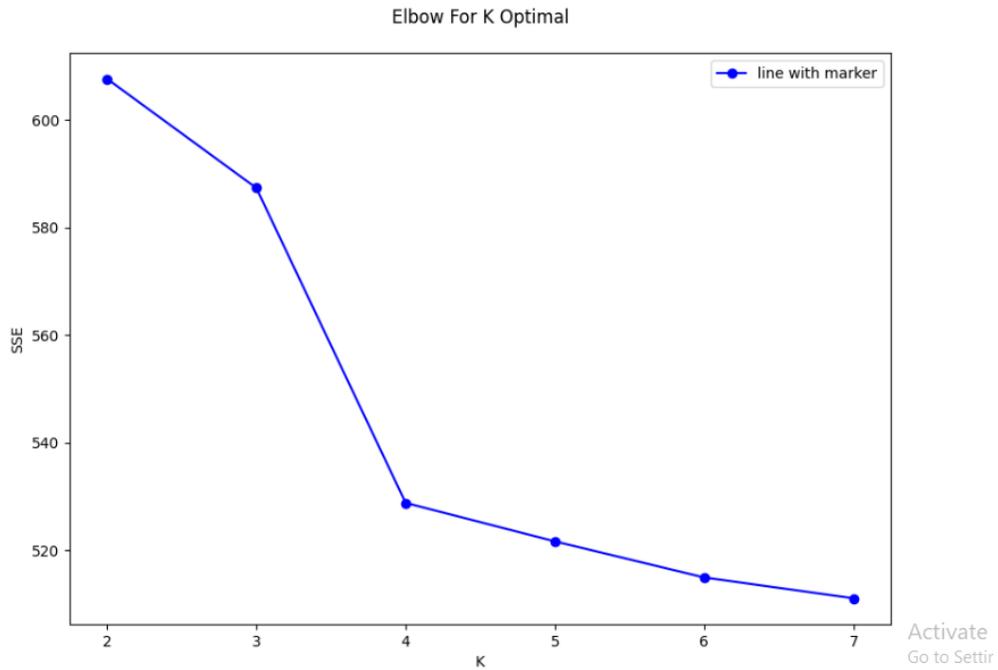
Table 2 Hasil Transformation dan Normalization

Jumlah Mahasiswa	Jumlah Dosen	Jumlah Program Studi	Fasilitas Olahraga	Wisma	Teknologi Informasi	Public Area	...	Akreditasi
1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	...	1.0000
0.5737	0.7726	0.7778	1.0000	1.0000	1.0000	1.0000	...	1.0000
0.8742	0.8376	0.5833	1.0000	1.0000	1.0000	1.0000	...	1.0000
0.3592	0.4594	0.0833	1.0000	1.0000	1.0000	1.0000	...	1.0000
0.3459	0.5220	0.7500	1.0000	1.0000	1.0000	1.0000	...	1.0000
0.1096	0.1044	0.0833	1.0000	0.0000	1.0000	0.0000	...	1.0000
0.1208	0.0046	0.0000	1.0000	1.0000	1.0000	1.0000	...	1.0000
0.0000	0.4571	0.0833	1.0000	0.0000	1.0000	0.0000	...	1.0000
0.1055	0.0000	0.0556	1.0000	0.0000	0.0000	0.0000	...	1.0000
0.3093	0.2088	0.3611	1.0000	1.0000	1.0000	1.0000	...	1.0000
...	...	...	...	...	...	...	...	...
0.1055	0.0000	0.0556	1.0000	0.0000	0.0000	0.0000	...	1.0000

Dari data diatas adalah hasil dari transformasi data, dimana dalam penelitian ini dataset yang digunakan memiliki atribut sebanyak 35 atribut yaitu Jumlah Mahasiswa, Jumlah Dosen, Jumlah Program Studi, Fasilitas Olahraga, Wisma/Asrama/Hotel, Teknologi Informasi, Public Area, PERS Mahasiswa, Asuransi, Gedung Pertemuan, Laboratorium, Poliklinik, Bus Kampus, Kalender Pendidikan, Pusat Pelatihan Bahasa, Perpustakaan, Sarana Ibadah, *Free Hotspot*, Jumlah Fasilitas, Bus Umum, Restaurant, Tempat Ibadah, Kos/Asrama, Terminal Bus, Café, Tempat Olahraga, Bandara, Mall, Rumah Sakit, Stasiun Kereta, Supermarket, Apotek, Jumlah Tempat Umum Terdekat, Akreditasi, dan Ranking Nasional. Dari 35 atribut tersebut terdapat atribut nominal, yaitu atribut akreditasi dibagi menjadi 3 golongan yaitu 1 menandakan akreditasi C, 2 menandakan akreditasi B, dan 3 menandakan akreditasi A. Setelah tahap transformasi dilakukan maka semua data yang ada dilakukan normalisasi untuk menyeimbangkan data agar semua dataset berada pada range 0 sampai dengan 1 menggunakan metode *scaling*

### B. Penentuan *Cluster*

Metode *Elbow* merupakan metode yang digunakan untuk menentukan jumlah *cluster* yang optimal dari data yang dimiliki yaitu data perguruan tinggi swasta, untuk menentukan jumlah *cluster* yang optimal dengan melihat selisih/*sum of square error* dari setiap *cluster*, jika *cluster* yang memiliki jumlah penurunan *sum of square* yang signifikan maka dapat ditentukan jumlah *cluster* yang optimal.



Gambar 2 Grafik Elbow

Gambar 2 merupakan hasil perhitungan *elbow* pada sistem didapat hasil K yang optimal yaitu 4 karena terlihat pada grafik *elbow* nilai sudut terkecil berada pada titik K=4.

### C. Hasil *Clustering*

Hasil *cluster* dapat dilihat setelah data yang telah dimasukkan kedalam sistem akan dilakukan proses klusterisasi. Proses klusterisasi dilakukan dengan menginputkan nilai K yang dapat dilihat pada gambar 3.

### Hitung Cluster

Jumlah Cluster

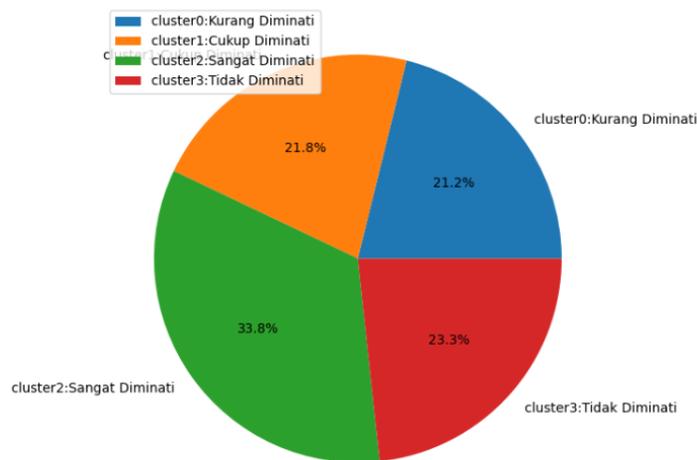
4

Hitung cluster!

Bandara	Mall	Rumah Sakit	Stasiun Kereta	Supermarket	Apotek	Jumlah Tempat Umum Terdekat	Akreditasi	Ranking Nasional	Cluster
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.000571265	2
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.001999429	2
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.002285061	2

Gambar 3 Hasil Cluster

Setelah menghitung *cluster* maka dapat di lihat hasil presentase dari setiap *cluster* yang dapat dilihat pada gambar 4.



Gambar 4 Persentase Setiap Cluster

#### D. Evaluasi Pengujian

pengujian *cluster* yang telah terbentuk dengan melihat nilai *silhouette* terbesar, yang dapat dilihat pada gambar 5.

Nilai K	Nilai Silhouette
4	0.7258934037794803

Gambar 5 Hasil Pengujian

Berdasarkan hasil perhitungan yang telah dilakukan seperti pada gambar 5 , dapat dilihat bahwa nilai *Silhouette* pada cluster 4 yaitu 0.7259 Dimana dengan nilai yang didapat tersebut tergolong kedalam *cluster Strong Structure*.

#### E. Interpretasi

Dari hasil *cluster* yang terbentuk yaitu 4 *cluster* dengan kategori sangat diminati, cukup diminati, kurang diminati, dan tidak diminati. Berikut penjelasan mengenai kategori untuk setiap *cluster* dan interpretasi analisis:

- *Cluster 2*, analisis terhadap *cluster* ini tergolong kedalam *cluster* sangat diminati karena dilihat dari data untuk setiap atribut memiliki nilai yang tinggi seperti Akreditasi yang memiliki nilai A, rata-rata memiliki fasilitas PERS Mahasiswa, memiliki lokasi yang strategis, dan Ranking Nasional terbaik jika dibandingkan dengan *cluster* lainnya.
- *Cluster 1*, analisis terhadap *cluster* ini tergolong kedalam *cluster* cukup diminati karena dilihat dari data untuk atribut Jumlah Program Studi , Jumlah Mahasiswa, Jumlah Dosen memiliki nilai tinggi, memiliki lokasi yang cukup strategis. Namun jika dilihat dari atribut Akreditasi, Jumlah Program Studi , Jumlah Mahasiswa, dan Jumlah Dosen memiliki nilai yang lebih kecil dari *cluster 2*.
- *Cluster 0*, analisis terhadap *cluster* ini tergolong kedalam *cluster* kurang diminati karena dilihat dari data untuk atribut Jumlah Fasilitas, Jumlah Tempat Umum Terdekat memiliki nilai yang kecil, terutama pada atribut Akreditasi rata-rata memiliki nilai C.
- *Cluster 3*, analisis terhadap *cluster* ini tergolong kedalam *cluster* tidak diminati karena dilihat dari setiap atribut yang ada memiliki nilai yang paling kecil dibandingkan dengan *cluster* lainnya, dan rata-rata tidak terakreditasi.

#### V. Kesimpulan

Dalam penelitian Klasterisasi Perguruan Tinggi Swasta Berdasarkan Minat Siswa Menggunakan Metode *K-Medoids* yang telah dilakukan dapat menghasilkan beberapa *cluster*, dengan penentuan jumlah *cluster* yang optimal menggunakan metode *Elbow*. Hasil yang diperoleh berdasarkan proses klasterisasi *K-Medoids* yang menghasilkan 4 buah *cluster* dimana *cluster 2* merupakan perguruan tinggi swasta yang sangat diminati, *cluster 1* merupakan perguruan tinggi swasta yang cukup diminati, *cluster 0* merupakan perguruan tinggi swasta yang kurang diminati, dan *cluster 3* merupakan perguruan tinggi swasta yang tidak diminati. Hasil nilai evaluasi dengan menggunakan metode *Silhouette Coefficient* untuk 4 *cluster* yaitu 0,72589 sehingga hasil *cluster* yang terbentuk masuk kedalam *cluster strong structure* atau merupakan *cluster* terbaik. Dari hasil evaluasi tersebut dapat ditarik kesimpulan bahwa algoritma *K-Medoids* sangat baik dalam mengolompokan perguruan tinggi swasta.

## REFERENCES

- [1] E. N. Wahyudi, A. Jananto, and Narwati, "Analisa Profil Data Mahasiswa Baru terhadap Program Studi yang dipilih di Perguruan Tinggi Swasta Jawa Tengah dengan Menggunakan Teknik Data Mining," *J. Teknol. Inf. Din.*, vol. 16, no. 1, pp. 29–43, 2011, [Online]. Available: <https://unisbank.ac.id/ojs/index.php/fti1/article/view/348%0Ahttps://unisbank.ac.id/ojs/index.php/fti1/article/view/348/225>.
- [2] W. S. D. E. Tikwalau, "Faktor-Faktor Yang Mempengaruhi Siswa-Siswa SMU/SMK Terhadap Keputusan Pemilihan Perguruan Tinggi Swasta," *Univ. Muhammadiyah Jakarta*, pp. 1–14, 2019.
- [3] H. Sawaji and & Taba, "An Analysis of Student's Decision Making to Choose Private Universities in South Sulawesi," *E-Journal Pascasarj. Univ. Hassanudin*, no. 4, pp. 1–22, 2011.
- [4] R. Rizalihadi, P. Poningsih, and ..., "Penerapan Algoritma K-Medoids Dalam Menentukan Jurusan Di SMA Swasta Kartika I-4," *Pros. ....*, vol. 1, pp. 683–691, 2019, [Online]. Available: <http://bulletin.indoms-acehsumut.org/index.php/simantap/article/view/221>.
- [5] P. Arora, Deepali, and S. Varshney, "Analysis of K-Means and K-Medoids Algorithm for Big Data," *Phys. Procedia*, vol. 78, no. December 2015, pp. 507–512, 2016, doi: 10.1016/j.procs.2016.02.095.
- [6] I. Gunawan, G. Anggraeni, E. S. Rini, and Y. Mustofa, "Klasterisasi provinsi di Indonesia berbasis perkembangan kasus Covid-19 menggunakan metode K-Medoids," *Semin. Nas. Mat. dan Pendidik. Mat.*, pp. 301–306, 2020.
- [7] A. Onan, "A K-medoids Based Clustering Scheme Z ith an \$ plication to ' ocument & lustering," 2017.
- [8] A. Rad, B. Naderi, and M. Soltani, "Clustering and ranking university majors using data mining and AHP algorithms: A case study in Iran," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 755–763, 2011, doi: 10.1016/j.eswa.2010.07.029.
- [9] R. Nainggolan, R. Perangin-Angin, E. Simarmata, and A. F. Tarigan, "Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method," *J. Phys. Conf. Ser.*, vol. 1361, no. 1, 2019, doi: 10.1088/1742-6596/1361/1/012015.
- [10] J. Zhou, Y. Pan, C. L. P. Chen, D. Wang, and S. Han, "K-medoids method based on divergence for uncertain data clustering," *2016 IEEE Int. Conf. Syst. Man, Cybern. SMC 2016 - Conf. Proc.*, pp. 2671–2674, 2017, doi: 10.1109/SMC.2016.7844643.
- [11] T. K. Barik and V. A. Centeno, "K-Medoids Clustering of Setting Groups in Directional Overcurrent Relays for Distribution System Protection," *2020 IEEE Kansas Power Energy Conf. KPEC 2020*, 2020, doi: 10.1109/KPEC47870.2020.9167531.
- [12] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.
- [13] H. B. Zhou and J. T. Gao, "Automatic method for determining cluster number based on silhouette coefficient," *Adv. Mater. Res.*, vol. 951, pp. 227–230, 2014, doi: 10.4028/www.scientific.net/AMR.951.227.
- [14] S. Luan, X. Kong, B. Wang, Y. Guo, and X. You, "Silhouette coefficient based approach on cell-phone classification for unknown source images," *IEEE Int. Conf. Commun.*, pp. 6744–6747, 2012, doi: 10.1109/ICC.2012.6364928.
- [15] R. Raksanagara, Y. H. Chrisnanto, and A. I. Hadiana, "Analisis Sentimen Jasa Ekspedisi Barang Menggunakan Metode Naïve Bayes," *Pros. SNST ke-7 Tahun 2016*, pp. 19–24, 2016.