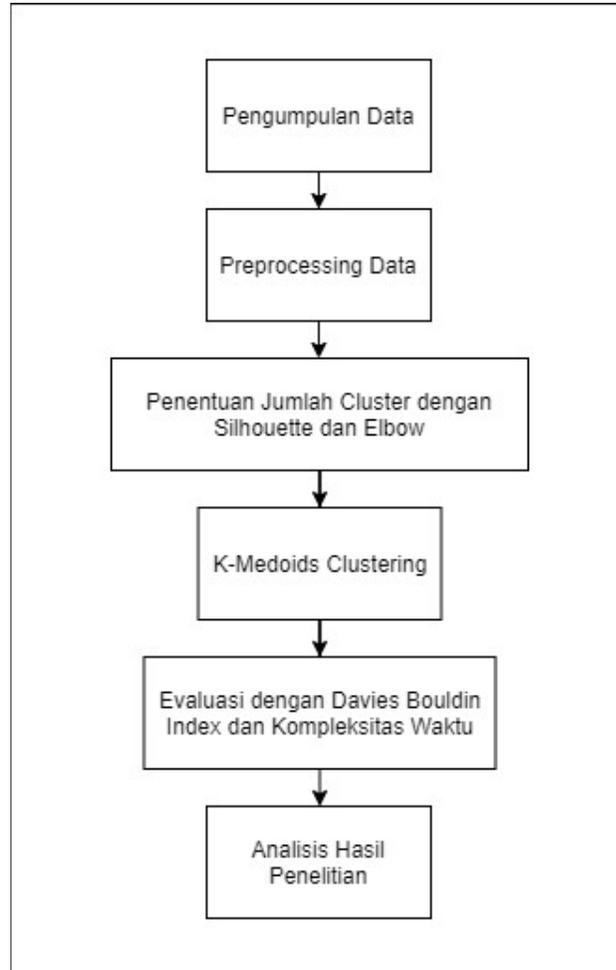


BAB III

ANALISIS DAN PERANCANGAN



Gambar 3.1 Tahapan dalam Penelitian

Pada Gambar 3.1 merupakan tahap yang akan dilakukan dalam penelitian ini yakni pengumpulan data, *Preprocessing data*, mencari jumlah *cluster* optimal dengan *Silhouette Coefficient* dan *Elbow Method*, lalu meng*Clustering* data dengan *K-Medoids*, evaluasi hasil *Clustering* dengan *Davies Bouldin Index* dan pengujian kompleksitas waktu lalu menganalisis hasil penelitian.

3.1 Pengumpulan Data

Pengumpulan data merupakan tahapan pertama yang dilakukan dalam pengerjaan tugas akhir ini. Proses pengumpulan data evaluasi kinerja dosen diperoleh dari SPM Institut Teknologi Sumatera yang mana sudah mengumpulkan kuisisioner evaluasi kinerja dosen dari mahasiswa yang dilakukan pada akhir semester. Untuk daftar pertanyaannya dapat dilihat pada Tabel 3.1, lalu untuk isi datanya dapat dilihat pada Tabel 3.2.

Tabel 3.1 Daftar Pertanyaan

Atribut	Pertanyaan
1	Dosen menyiapkan materi dengan sebaik-baiknya untuk diajarkan.
2	Dosen memberikan informasi yang cukup tentang hal-hal tertentu yang harus Saya capai atau kuasai (luaran mata kuliah) setelah mengikuti mata kuliah ini termasuk materi apa yang akan dipelajari serta cara penilaian yang dilakukan.
3	Dosen memakai buku acuan/diktat/ bahan ajar yang mana memang membantu mahasiswa dalam menguasai materi pelajaran.
4	Beban kerja mata kuliah ini sesuai dengan SKS-nya.
5	Dosen memberikan penilaian berdasarkan evaluasi yang lebih dari 1 kali (ujian/tugas/pekerjaan rumah) yang relevan dengan luaran mata kuliah.
6	Dosen menguasai materi perkuliahan dengan sangat baik.
7	Dosen berkomunikasi dengan efektif dan berusaha dengan caranya agar mahasiswa tertarik mengikuti pelajarannya.
8	Dosen bersifat responsif dan bersedia berdiskusi terkait mata kuliah di dalam ataupun di luar jam kuliah.
9	Dosen menggunakan waktu kuliah dengan semestinya.
10	Apabila berhalangan hadir, dosen memberikan waktu pengganti kepada mahasiswa.
11	Dosen berlaku adil kepada seluruh mahasiswa tanpa memandang suku, agama, ras, dan golongan

Atribut	Pertanyaan
12	Sarana dan prasarana pembelajaran mata kuliah tersedia dengan sangat baik (kesediaan ruangan berikut papan tulis, spidol, proyektor, dan sebagainya)
13	Anda mencapai atau menguasai luaran dari mata kuliah ini
14	Tingkat kehadiran Saya dalam mata kuliah ini tinggi
15	Forum tanya jawab/diskusi perkuliahan berjalan dengan baik dimana mahasiswa dapat memberikan pertanyaan atau umpan balik dan dosen memberikan tanggapan
16	Penyelenggaraan kuliah dengan menggunakan metode e-learning secara keseluruhan berjalan dengan efektif

Tabel 3.2 Isi Data

No	Pertanyaan	Sangat Setuju	Setuju	Tidak setuju	Sangat tidak setuju
1	Dosen menyiapkan materi dengan sebaik-baiknya untuk diajarkan	62.00 %	32.00 %	3.00 %	3.00 %
2	Dosen memberikan informasi yang cukup tentang hal-hal tertentu yang harus Saya capai atau kuasai (luaran mata kuliah) setelah mengikuti mata kuliah ini termasuk materi apa yang akan dipelajari serta cara penilaian yang dilakukan	47.00 %	52.00 %	0.00 %	2.00 %
3	Dosen memakai buku acuan/diktat/ bahan ajar yang mana memang membantu mahasiswa dalam menguasai materi pelajaran	53.00 %	47.00 %	0.00 %	0.00 %
4	Beban kerja mata kuliah ini sesuai dengan SKS-nya	50.00 %	45.00 %	5.00 %	0.00 %
5	Dosen memberikan penilaian berdasarkan evaluasi yang lebih dari 1 kali (ujian/tugas/pekerjaan rumah) yang relevan dengan luaran mata kuliah	53.00 %	43.00 %	3.00 %	0.00 %
6	Dosen menguasai materi perkuliahan dengan sangat baik	52.00 %	47.00 %	2.00 %	0.00 %
7	Dosen berkomunikasi dengan efektif dan berusaha dengan caranya agar mahasiswa tertarik mengikuti pelajarannya	52.00 %	45.00 %	3.00 %	0.00 %
8	Dosen bersifat responsif dan bersedia berdiskusi terkait mata kuliah di dalam ataupun di luar jam kuliah	53.00 %	42.00 %	3.00 %	2.00 %
9	Dosen menggunakan waktu kuliah dengan semestinya	53.00 %	35.00 %	5.00 %	7.00 %
10	Apabila berhalangan hadir, dosen memberikan waktu pengganti kepada mahasiswa	55.00 %	38.00 %	5.00 %	2.00 %
11	Dosen berlaku adil kepada seluruh mahasiswa tanpa memandang suku, agama, ras, dan golongan	58.00 %	38.00 %	3.00 %	0.00 %
12	Sarana dan prasarana pembelajaran mata kuliah tersedia dengan sangat baik (kesediaan ruangan berikut papan tulis, spidol, proyektor, dan sebagainya)	45.00 %	48.00 %	5.00 %	2.00 %

No	Pertanyaan	Sangat Setuju	Setuju	Tidak setuju	Sangat tidak setuju
13	Anda mencapai atau menguasai luaran dari mata kuliah ini	35.00 %	58.00 %	5.00 %	2.00 %
14	Tingkat kehadiran Saya dalam mata kuliah ini tinggi	63.00 %	33.00 %	3.00 %	0.00 %
15	Forum tanya jawab/diskusi perkuliahan berjalan dengan baik dimana mahasiswa dapat memberikan pertanyaan atau umpan balik dan dosen memberikan tanggapan	52.00 %	45.00 %	3.00 %	0.00 %
16	Penyelenggaraan kuliah dengan menggunakan metode e-learning secara keseluruhan berjalan dengan efektif	50.00 %	43.00 %	5.00 %	2.00 %

3.2 Preprocessing data

Preprocessing adalah proses pembersihan data yang dilakukan sebelum ekstraksi fitur. *Preprocessing* dibutuhkan untuk mengurangi data yang tidak dibutuhkan, misalkan penghapusan atribut. Dalam *Preprocessing data* ini akan dilakukan 3 tahap yaitu *Data Transformation*, *Data Selection* dan *Data Integration*.

4.2.1 Data Transformation

Transformasi data merupakan tahap dilakukannya perubahan data atau konversi data, seperti perubahan dari tipe nominal ke tipe kategori ataupun sebaliknya. Pada penelitian ini dilakukan proses transformasi data dari persen menjadi desimal. Setiap nilai atribut akan dikalikan dengan bobot atributnya, bobot dari atribut dapat dilihat pada Tabel 3.4, lalu akan menambahkan atribut baru yaitu “Jumlah”.

Tabel 3.3 Bobot Atribut

Nama atribut	Bobot
Sangat setuju	4
Setuju	3
Tidak setuju	2
Sangat tidak setuju	1

4.2.2 Data Selection

Data Selection adalah Proses seleksi atau pemilihan data sesuai kebutuhan terhadap analisis untuk diterima dari koleksi data yang ada. Pada

penelitian kali ini atribut yang dibutuhkan yaitu atribut pertanyaan dan jumlah saja.

4.2.3 Data Integration

Data Integration adalah proses di mana data dari banyak sumber masuk ke satu lokasi. Pada tahap ini akan melakukan penggabungan dari hasil *Data Selection* pada seluruh dosen yang akan dimasukkan kedalam satu file.

3.3 Mencari Jumlah Cluster Optimal

Melakukan pencarian jumlah *cluster* teroptimal dengan silhouette dan *Elbow* yang akan digunakan untuk melakukan *Clustering* guna optimalisasi metode *K-Medoids*.

3.3.1 Silhouette Coefficient

Tahapan-tahapan menentukan nilai k teroptimal pada *K-Medoids Clustering* dengan metode *Silhouette Coefficient* sebagai berikut :

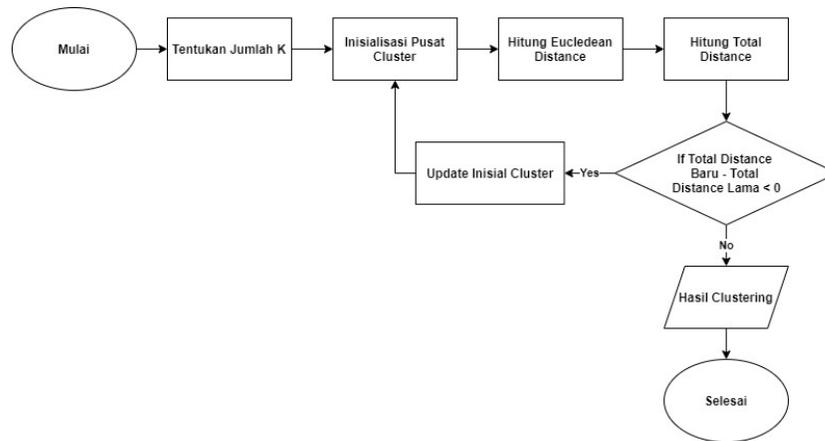
1. Lakukan *Clustering* algoritma *K-Medoids* pada setiap nilai k, mulai dari k=2 hingga 6.
2. Hitung hasil *Silhouette Coefficient*.
3. Melihat hasil *Silhouette Coefficient* dari nilai k yang memiliki *score* tertinggi.
4. Tetapkan k dengan *score Silhouette Coefficient* tertinggi sebagai jumlah *cluster* teroptimal.

3.3.2 Elbow Method

Tahapan-tahapan menentukan nilai k teroptimal pada *K-Medoids Clustering* dengan metode *Elbow* sebagai berikut :

1. Lakukan *Clustering* algoritma *K-Medoids* pada setiap nilai k, mulai dari k=1 hingga 6.
2. Hitung hasil sum of square error dari tiap nilai K.
3. Melihat hasil sum of square error dari nilai K yang turun secara drastis atau membentuk siku.
4. Tetapkan nilai K yang berbentuk siku sebagai jumlah *cluster* teroptimal.

3.4 K-Medoids Clustering



Gambar 3.2 Flowchart *K-Medoids*

Proses *Clustering* menggunakan metode *K-Medoids* yang diterapkan pada penelitian ini dapat diamati pada Gambar 3.2 [5].

Proses yang dilakukan dalam penerapan metode ini adalah sebagai berikut:

1. Tentukan jumlah *cluster*, *cluster* yang akan dibentuk yaitu sebanyak 2 *cluster*.
2. Inisialisasi pusat *cluster* sebanyak k (jumlah *cluster*). Tabel 3.4 merupakan tabel *Centroid Awal K-Medoids* yang mana *centroid* awal berada pada “Nama = A” untuk *centroid* 1 dan “Nama = G” untuk *centroid* 2.

Tabel 3.4 *Centroid Awal K-Medoids*

Nama	A1	A2	Medoids
A	3.53	3.46	1
B	3.57	3.53	
C	3.5	3.5	
D	3.52	3.42	
E	3.35	3.28	
F	3.72	3.66	
G	3.2	3.18	2
H	3.19	3.17	
I	3.32	3.27	
J	3.53	3.4	

3. Hitung jarak masing-masing objek ke medoid sementara dengan menggunakan Euclidean *Distance*, kemudian tandai jarak terdekat objek ke medoid dan hitung totalnya.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3.1)$$

Jarak medoid 1 (3.53, 3.46) ke data 1 (3.53, 3.46) =

$$\begin{aligned} (3.53, 3.46) \leftrightarrow (3.53, 3.46) &= \sqrt{|3.53 - 3.53|^2 + |3.46 - 3.46|^2} \\ &= 0 \end{aligned}$$

Kemudian hitung jarak medoid 2 (3.2, 3.18) ke data 1 (3.53, 3.46) =

$$\begin{aligned} ((3.2, 3.18) \leftrightarrow (3.53, 3.46)) &= \sqrt{|3.2 - 3.53|^2 + |3.18 - 3.46|^2} \\ &= \sqrt{0.33^2 + 0.28^2} \\ &= 0.432782 \end{aligned}$$

Lakukan perhitungan ke semua data, dan menghasilkan nilai seperti pada Tabel 3.5. Tabel 3.5 merupakan tabel hasil *cluster* medoid awal yang berisikan atribut “Nama” mempresentasikan nama dosen, “A1” mempresentasikan *score* pertanyaan 1, “A2” mempresentasikan *score* pertanyaan 2, “Jarak ke medoid 1” mempresentasikan jarak data ke medoid 1, “Jarak ke medoid 2” mempresentasikan jarak data ke medoid 2, “kedekatan” mempresentasikan jarak minimum antaran jarak ke medoid 1 dan jarak ke medoid 2 dan “*cluster*” mempresentasikan dosen tersebut masuk ke *cluster* yang mana

Tabel 3.5 Hasil *Cluster* Medoid Awal

Nama	A1	A2	Jarak ke medoid 1	Jarak ke medoid 2	kedekatan	<i>cluster</i>
A	3.53	3.46	0	0.432782	0	0
B	3.57	3.53	0.080623	0.509313	0.080623	0
C	3.5	3.5	0.05	0.438634	0.05	0
D	3.52	3.42	0.041231	0.4	0.041231	0
E	3.35	3.28	0.254558	0.180278	0.180278	1
F	3.72	3.66	0.275862	0.707672	0.275862	0
G	3.2	3.18	0.432782	0	0	1

Nama	A1	A2	Jarak ke medoid 1	Jarak ke medoid 2	kedekatan	cluster
H	3.19	3.17	0.446878	0.014142	0.014142	1
I	3.32	3.27	0.283196	0.15	0.15	1
J	3.53	3.4	0.06	0.396611	0.06	0
Total Distance					0.852135618	

4. Pilih secara acak objek pada masing-masing *cluster* sebagai kandidat untuk medoid baru. Pilih data 1 (100, 50) sebagai medoid-1 dan data 5 (65, 40) sebagai medoid-2. Tabel *centroid* baru *k-medoids* dapat dilihat pada Tabel 3.6.

Tabel 3.6 *Centroid* Baru *K-Medoids*

Nama	A1	A2	Medoids
A	3.53	3.46	
B	3.57	3.53	
C	3.5	3.5	
D	3.52	3.42	
E	3.35	3.28	
F	3.72	3.66	
G	3.2	3.18	
H	3.19	3.17	
I	3.32	3.27	1
J	3.53	3.4	2

5. Hitung jarak setiap objek yang berada pada masing-masing *cluster* dengan kandidat medoid baru.

Jarak medoid 1 (3.32, 3.27) ke data 1 (3.53, 3.46) =

$$\begin{aligned}
 (3.32, 3.27) \leftrightarrow (3.53, 3.46) &= \sqrt{|3.32 - 3.53|^2 + |3.27 - 3.46|^2} \\
 &= \sqrt{0.21^2 + 0.19^2} \\
 &= 0.283196
 \end{aligned}$$

Kemudian hitung jarak medoid 2 (3.53, 3.4) ke data 1 (3.53, 3.46) =

$$((3.53, 3.4) \leftrightarrow (3.53, 3.46)) = \sqrt{|3.53 - 3.53|^2 + |3.4 - 3.46|^2}$$

$$= \sqrt{0 + 0.06^2}$$

$$= 0.06$$

Lakukan perhitungan ke semua data, dan menghasilkan nilai seperti pada Tabel 3.7. Tabel 3.7 merupakan tabel hasil *cluster* medoid baru yang mempunyai atribut “Nama” mempresentasikan nama dosen, “A1” mempresentasikan *score* pertanyaan 1, “A2” mempresentasikan *score* pertanyaan 2, “Jarak ke medoid 1” mempresentasikan jarak data ke medoid 1, “Jarak ke medoid 2” mempresentasikan jarak data ke medoid 2, “kedekatan” mempresentasikan jarak minimum antaran jarak ke medoid 1 dan jarak ke medoid 2 dan “*cluster*” mempresentasikan dosen tersebut masuk ke *cluster* yang mana.

Tabel 3.7 Hasil *Cluster* Medoid Baru

Nama	A1	A2	Jarak ke medoid 1	Jarak ke medoid 2	Kedekatan	<i>Cluster</i>
A	3.53	3.46	0.283196	0.06	0.06	1
B	3.57	3.53	0.360694	0.136015	0.136015	1
C	3.5	3.5	0.292062	0.104403	0.104403	1
D	3.52	3.42	0.25	0.022361	0.022361	1
E	3.35	3.28	0.031623	0.216333	0.031623	0
F	3.72	3.66	0.558659	0.322025	0.322025	1
G	3.2	3.18	0.15	0.396611	0.15	0
H	3.19	3.17	0.164012	0.410488	0.164012	0
I	3.32	3.27	0	0.246982	0	0
J	3.53	3.4	0.246982	0	0	1
Total <i>Distance</i>					0.990438265	

6. Hitung total simpangan (S) Jika a adalah jumlah jarak terdekat antara objek ke medoid awal, dan b adalah jumlah jarak terdekat antara objek ke medoid baru, maka total simpangan adalah $S = b - a$ Jika $S < 0$, maka kembali ke langkah 4 dan hentikan jika $S > 0$.

$$S = \text{total } distance \text{ baru} - \text{total } distance \text{ lama}$$

$$= 0.990438265 - 0.852135618$$

$$= 0.138302647$$

3.5 Evaluasi Hasil *Clustering*

3.5.1 Perhitungan Davies Bouldin Index

Hasil *cluster* yang didapatkan dari proses *Clustering* dengan *K-Medoids* ini dianalisis dengan menggunakan metode *Davies Bouldin Index*. Pada pengukuran kualitas *Clustering* dengan menggunakan metode *Davies Bouldin Index* semakin kecil nilainya akan semakin baik kualitas *clusternya*, Tahapan-tahapan metode ini sebagai berikut:

1. Menentukan *centroid* pada tiap *cluster*.

$$\text{Centroid 0} = \text{atribut1} = 3.5 + 3.57 + 3.5 + 3.52 + 3.72 + 3.53 = 3.561667$$

$$\text{atribut2} = 3.46 + 3.53 + 3.5 + 3.42 + 3.66 + 3.4 = 3.495$$

$$\text{Centroid 1} = \text{atribut1} = 3.35 + 3.32 + 3.19 + 3.32 = 3.265$$

$$\text{atribut2} = 3.28 + 3.28 + 3.17 + 3.27 = 3.225$$

2. Menghitung nilai *Sum of square within cluster* (SSW)

$$SSWi = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \quad (3.2)$$

Hitung jarak data ke tiap *cluster* dengan rumus *Euclidean Distance*.

Jarak *centroid 0* (3.561667, 3.495) ke data *cluster 0* pertama (3.53, 3.46)

$$= (3.561667, 3.495) \leftrightarrow (3.53, 3.46)$$

$$= \sqrt{(3.561667 - 3.53)^2 + (3.495 - 3.46)^2}$$

$$= 0.047199$$

Kemudian hitung jarak *centroid 1* (3.265, 3.225) ke data *cluster 1* pertama (3.35, 3.28) =

$$(3.265, 3.225) \leftrightarrow (3.35, 3.28) = \sqrt{(3.265 - 3.35)^2 + (3.225 - 3.28)^2}$$

$$= 0.101242$$

Lakukan perhitungan ke semua data pada tiap *cluster* lalu rata-rata kan jarak pada tiap *cluster*, dan menghasilkan nilai seperti pada Tabel 3.8. Tabel 3.8. merupakan tabel hasil SSW yang mempunyai atribut “jarak” yang merepresentasikan jarak data ke masing-masing *cluster* dan SSW pada setiap *cluster*.

Tabel 3.8 Hasil SSW

Nama	A1	A2	Cluster	Jarak
A	3.53	3.46	0	0.047199
B	3.57	3.53	0	0.035978
C	3.5	3.5	0	0.061869
D	3.52	3.42	0	0.085797
F	3.72	3.66	0	0.22868
J	3.53	3.4	0	0.100139
SSW Cluster 0				0.093277
E	3.35	3.28	1	0.101242
G	3.2	3.18	1	0.079057
H	3.19	3.17	1	0.093005
I	3.32	3.27	1	0.071063
SSW Cluster 1				0.086092

3. Menghitung nilai *Sum of square between cluster* (SSB)

$$SSB_{i,j} = d(c_i, c_j) \quad (3.3)$$

Hitung jarak *centroid* 0 ke *centroid* 1 dengan rumus *Euclidean Distance*.

Jarak *centroid* 0 (3.561667, 3.495) ke *centroid* 1 (3.265, 3.225)

$$SSB = (3.561667, 3.495) \leftrightarrow (3.265, 3.225)$$

$$SSB = \sqrt{(3.561667 - 3.265)^2 + (3.495 - 3.225)^2}$$

$$SSB = 0.401137$$

4. Melakukan pengukuran rasio (R_{ij})

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{ij}} \quad (3.4)$$

$$= \frac{0.093277 + 0.086092}{0.401137}$$

$$= 0.447151$$

Karna hanya ada dua *cluster* hasil rasio hanya satu.

5. Menghitung nilai *Davies Bouldin Index* (DBI). Dengan rumus sebagai berikut:

$$DBI = \frac{1}{K} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (3.5)$$

Jadi nilai DBI nya adalah :

$$DBI = \frac{0.447151 + 0.447151}{2}$$

$$DBI = 0.447151$$

3.5.2 Pengujian Kompleksitas Waktu

Tahapan-tahapan pengujian kompleksitas waktu pada setiap metode *Silhouette Coefficient* dan *Elbow* berdasarkan waktu yang dibutuhkan untuk memperoleh hasil jumlah *cluster* optimal sebagai berikut :

1. Lakukan pencarian *cluster* optimal pada metode *Silhouette Coefficient* dan *Elbow* sebanyak 30 kali percobaan dan catat waktu setiap percobaannya.
2. Hasil dari 30 kali percobaan tersebut maka hitunglah rata-ratanya. Waktu rata-rata tersebut merupakan waktu yang menjadi hasil pengujian kompleksitas waktu untuk setiap hasil pencarian *cluster* optimal.