**BAB IV**

**HASIL DAN PEMBAHASAN**

* 1. **Hasil Pengumpulan Data**

Data yang digunakan dalam penelitian ini dibatasi dari bulan Maret 2018 hingga bulan Agustus 2018 dengan jumlah data 4325. Kemudian yang menjadi fokus dalam penelitian ini adalah melakukan pengelompokan pada data pelanggan agen Lion Parcel dengan menggunakan algoritma Fuzzy C-Means.

Data berupa file exel dengan file “data.xls”. Di dalam file tersebut terdapat sebuat sheet dengan nama “ Sheet 1”. Dalam Sheet 1 terdapat sebuah tabel dengan 4325 baris dan 13 kolom (atribut) yaitu: SNo, BookingDate, ID, STTNo, Shipper, OriginCity, DestinationCity, ForwardArea, Product, GrossWt, VolumeWt, CommodityName, LastTrackingStatus dapat dilihat pada Tabel 4.1.

Tabel 4.1 Data Pelanggan Agen Lion Parcel

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SNo | BookingDate | ID | STTNo | Shipper | OriginCity | DestinationCity | ForwardArea | Product | GrossWt | VolumeWt | CommodityName | LastTrackingStatus |
| 110459075 | 01-Mar-2018 | JOG-1932 | 10-17-1617709 | DEE | JOGJAKARTA | GORONTALO |  | REGPACK | 7 | 3 | PRODUCT COSMETICS NON LIQUID | RECEIPT BY RIA . |
| 110459088 | 01-Mar-2018 | JOG-1932 | 10-17-1617711 | DEE | JOGJAKARTA | GRESIK |  | REGPACK | 1 | 1 | PRODUCT COSMETICS NON LIQUID | RECEIPT BY NURFAN . |
| 110459140 | 01-Mar-2018 | JOG-1932 | 10-17-1617647 | MEC ARTUR | JOGJAKARTA | MANOKWARI |  | REGPACK | 2 | 1 | ACCESSORIES | RECEIPT BY ANDRY TRIAS . |
| 110459171 | 01-Mar-2018 | JOG-1932 | 11-18-315622 | MELVIN | JOGJAKARTA | AMBON |  | ONEPACK | 1 | 1 | ACCESSORIES | RECEIPT BY SANCHIA AHULUHELUW . |
| 110459198 | 01-Mar-2018 | JOG-1932 | 10-17-1617708 | VINDY SHOP | JOGJAKARTA | KENDARI |  | REGPACK | 1 | 1 | PRODUCT COSMETICS NON LIQUID | RECEIPT BY HESTI . |
| 110459247 | 01-Mar-2018 | JOG-1932 | 11-18-315695 | DEE | JOGJAKARTA | BATAM |  | REGPACK | 1 | 1 | PRODUCT COSMETICS NON LIQUID | RECEIPT BY IBU WIWIT . |
| 110459865 | 01-Mar-2018 | JOG-1932 | 11-18-316262 | EVIE SADDIYAH | JOGJAKARTA | NABIRE |  | REGPACK | 1 | 1 | ACCESSORIES | RECEIPT BY JUMAINI . |
| 110460042 | 01-Mar-2018 | JOG-1932 | 11-18-316433 | RIRIN AFRIANI | JOGJAKARTA | PANGKAL PINANG |  | REGPACK | 1 | 1 | PRODUCT COSMETICS NON LIQUID | RECEIPT BY ARIEL . |
| 110461292 | 01-Mar-2018 | JOG-1932 | 10-17-1618143 | VIVI MARLIANA | JOGJAKARTA | SAMARINDA |  | REGPACK | 22 | 17 | CLOTHING/GARMENT | RECEIPT BY IFAH . |

* 1. **Tahapan Pengembangan Sistem**
     1. Tahapan Data Mining

1. Selection Data

Tahapan seleksi dilakukan pemilihan data yang berupa variabel yang dibutuhkan dari data pelanggan yaitu SNo, Product, GrossWt dan VolumeWt. Hasil dari seleksi data pelanggan dapat dilihat Pada Tabel 4.2.

Tabel 4.2. : Tabel hasil Selection pada data pelanggan

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | SNo | Product | GrossWt | VolumeWt |
| 1 | 110459075 | REGPACK | 7 | 3 |
| 2 | 110459088 | REGPACK | 1 | 1 |
| 3 | 110459140 | REGPACK | 2 | 1 |
| 4 | 110459171 | ONEPACK | 1 | 1 |
| 5 | 110459198 | REGPACK | 1 | 1 |
| 6 | 110459247 | REGPACK | 1 | 1 |
| 7 | 110459865 | REGPACK | 1 | 1 |
| 8 | 110460042 | REGPACK | 1 | 1 |
| 9 | 110461292 | REGPACK | 22 | 17 |
| 10 | 110459075 | REGPACK | 7 | 3 |

* 1. **Tahapan Pengelompokkan K-Medois**

Dari algoritma Fuzzy C-Means metode Clustering maka dapat dilakukan perhitungan sebagai berikut:

1. Data yang dihitung adalah data yang telah di proses seleksi data.
2. Menentukan jumlah cluster data pelanggan yang akan dijadikan kelompok, data pelanggan ini sudah ditentukan menjadi 4 kelompok.
3. Variabel yang akan digunakan yaitu SNo, Product, GrossWt dan VolumeWt.

Setelah Proses *Eulidean Distance* digunakan untuk menghitung jarak terdekat pada setiap data ispa. Dikatakan jarak terdekat yaitu nilai jarak Euclidean yang paling kecil atau nilai medekti nol. Hasil proses Euclidean Distance dapat dilihat pada tabel 4.3:

Tabel 4.3 Data yang akan digunakan untuk menghitung jarak Eucludien Distance

|  |  |  |  |
| --- | --- | --- | --- |
|  | Berat Badan | Kondisi Lingkungan | Status gizi |
| D1 | 3.4 kg | 10 | 25.0 |
| D2 | 5.1 kg | 10 | 20.0 |
| D3 | 4.5 kg | 5 | 19.5 |
| D4 | 5.6 kg | 10 | 25.0 |
| D5 | 6.3 kg | 3 | 17.5 |
| D6 | 5.5 kg | 10 | 18.5 |
| D7 | 7.2 kg | 1 | 16.5 |
| D8 | 7.0 kg | 10 | 25.0 |
| D9 | 6.9 kg | 10 | 21.0 |
| D10 | 4.0 kg | 10 | 22.0 |

Rumus :

*de* = 2

= 5.28

Tabel 4.4 Tabel Hasil Euclidean Distance

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | D9 | D10 |
| D1 | 0 | 5.28 | 7.51 | 2.2 | 10.66 | 6.83 | 12.94 | 3.6 | 5.31 | 3.05 |
| D2 | 5.28 | 0 | 5.06 | 5.02 | 7.52 | 1.55 | 9.88 | 5.34 | 2.05 | 2.28 |
| D3 | 7.51 | 5.06 | 0 | 7.51 | 3.35 | 5.19 | 5.68 | 7.84 | 5.74 | 5.61 |
| D4 | 2.2 | 5.02 | 7.51 | 0 | 10.28 | 6.50 | 12.48 | 1.4 | 4.20 | 3.4 |
| D5 | 10.66 | 7.52 | 3.35 | 10.28 | 0 | 7.11 | 2.41 | 10.28 | 7.84 | 8.63 |
| D6 | 6.83 | 1.55 | 5.19 | 6.50 | 7.11 | 0 | 9.37 | 6.67 | 2.86 | 3.80 |
| D7 | 12.66 | 9.88 | 5.68 | 12.48 | 2.41 | 9.37 | 0 | 12.38 | 10.06 | 11.0 |
| D8 | 3.6 | 5.34 | 7.84 | 1.4 | 10.28 | 6.67 | 12.38 | 0 | 4.00 | 4.24 |
| D9 | 5.31 | 2.05 | 5.74 | 4.20 | 7.84 | 2.86 | 10.06 | 4.00 | 0 | 3.06 |
| D10 | 3.05 | 2.28 | 5.61 | 3.4 | 8.63 | 2.91 | 11.02 | 4.24 | 3.06 | 0 |

* + 1. Tahapan Pengelompokkan K--Medoids

Data percobaan pertama dengan 2 cluster yang menjadi titik pusat yaitu D2 dan D4. Dengan menggunakan metode k-medoids untuk menentukan keanggotaan seluruh dokumen terhadap medoids D2 dan D4. Bandingkan seluruh data dengan nilai Euclidean D2 dan D4, jika jarak terhadap D2 lebih besar dari pada jarak terhadap D4 maka masuk dalam cluster 1 (medoids D2), jika sebaliknya maka masuk dalam cluster 2 (medoids D4) ditunjukkan pada Tabel 4.5 dan 4.6.

Tabel 4.5 Medoids tahap 1

|  |  |  |  |
| --- | --- | --- | --- |
|  | D2 | D4 | Cluster |
| D1 | 5.28 | 2.2 | 2 |
| D3 | 5.06 | 7.51 | 1 |
| D5 | 7.52 | 10.28 | 1 |
| D6 | 1.55 | 6.50 | 1 |
| D7 | 9.88 | 12.48 | 1 |
| D8 | 5.34 | 1.4 | 2 |
| D9 | 2.05 | 4.20 | 1 |
| D10 | 2.28 | 3.4 | 1 |

Tabel 4.6. Cluster Pertama

|  |  |
| --- | --- |
| Cluster 1 (titik D2) | Cluster 2 (titik D4) |
| D3 | D1 |
| D5 | D8 |
| D6 |  |
| D7 |  |
| D9 |  |
| D10 |  |

Proses langkah selanjutnya k-medoids akan membandingkan titik pusat pada cluster 1 terlebih dahulu dengan anggota pada cluster 1 ditunjukkan pada tabel 4.7.

Tabel 4.7 hasil 1 cluster 1

|  |
| --- |
| Cluster 1 |
| Titik pusat D2 |
| D3 |
| D5 |
| D6 |
| D7 |
| D9 |
| D10 |

Cluster 1 iterasi 1 : Jumlah jarak data yang berada pada medoids ke-1 dan ke-2, jika jumlah jarak medoid ke-1 lebih besar dari pada jumlah medoid ke-2 maka medoid pada cluster menjadi data ke-2, namun jika sebaliknya jumlah pada medoid ke-1 lebih kecil dari pada jumlah ke-2 maka medoid tetap data ke-1 ditunjukkan pada tabel 4.8.

Tabel 4.8 Cluster 1 iterasi 1

|  |  |  |
| --- | --- | --- |
|  | D2 | D3 |
| D1 | 5.28 | 7.51 |
| D2 | 0 | 5.06 |
| D3 | 5.06 | 0 |
| D4 | 5.02 | 7.51 |
| D5 | 7.52 | 3.35 |
| D6 | 1.55 | 5.19 |
| D7 | 9.88 | 5.68 |
| D8 | 5.34 | 7.84 |
| D9 | 2.05 | 5.74 |
| D10 | 2.28 | 5.61 |
| Jumlah | 43.98 | 53.49 |

Perhitungan medoid cluster 1 iterasi 1 =D2<D3, maka medoid tetap. Kemudian bandingkan dengan data selanjutnya yaitu D5 ditunjukkan pada tabel 4.9.

Tabel 4.9 Cluster 1 iterasi 2

|  |  |  |
| --- | --- | --- |
|  | D2 | D5 |
| D1 | 5.28 | 10.46 |
| D2 | 0 | 7.52 |
| D3 | 5.06 | 3.35 |
| D4 | 5.02 | 10.28 |
| D5 | 7.52 | 0 |
| D6 | 1.55 | 7.11 |
| D7 | 9.88 | 2.41 |
| D8 | 5.34 | 10.28 |
| D9 | 2.05 | 7.84 |
| D10 | 2.28 | 8.63 |
| Jumlah | 43.98 | 68.08 |

Perhitungan medoid cluster 1 iterasi 2 = D2<D5 maka medoid tetap. Kemudian bandingkan dengan selanjutnya yaitu D6 ditunjukkan pada tabel 4.10.

Tabel 4.10 Cluster 1 iterasi 3

|  |  |  |
| --- | --- | --- |
|  | D2 | D6 |
| D1 | 5.28 | 6.83 |
| D2 | 0 | 1.55 |
| D3 | 5.06 | 5.19 |
| D4 | 5.02 | 6.50 |
| D5 | 7.52 | 7.11 |
| D6 | 1.55 | 0 |
| D7 | 9.88 | 9.37 |
| D8 | 5.34 | 6.67 |
| D9 | 2.05 | 2.86 |
| D10 | 2.28 | 2.91 |
| Jumlah | 43.98 | 48.99 |

Perhitungan medoid cluster 1 iterasi 4 = D2<D6 maka medoid tetap. Kemudian bandingkan dengan selanjutnya yaitu D7 ditunjukkan pada tabel 4.11

Tabel 4.11 Cluster 1 iterasi 4

|  |  |  |
| --- | --- | --- |
|  | D2 | D7 |
| D1 | 5.28 | 12.94 |
| D2 | 0 | 9.88 |
| D3 | 5.06 | 5.68 |
| D4 | 5.02 | 12.48 |
| D5 | 7.52 | 2.41 |
| D6 | 1.55 | 9.37 |
| D7 | 9.88 | 0 |
| D8 | 5.34 | 12.38 |
| D9 | 2.05 | 10.06 |
| D10 | 2.28 | 11.02 |
| Jumlah | 43.98 | 86.22 |

Perhitungan medoid cluster 1 iterasi 4 = D2<D7 maka medoid tetap. Kemudian bandingkan dengan selanjutnya yaitu D9 ditunjukkan pada tabel 4.12

Tabel 4.12 Cluster 1 iterasi 5

|  |  |  |
| --- | --- | --- |
|  | D2 | D9 |
| D1 | 5.28 | 5.31 |
| D2 | 0 | 2.05 |
| D3 | 5.06 | 5.74 |
| D4 | 5.02 | 4.20 |
| D5 | 7.52 | 7.84 |
| D6 | 1.55 | 2.86 |
| D7 | 9.88 | 10.06 |
| D8 | 5.34 | 4.00 |
| D9 | 2.05 | 0 |
| D10 | 2.28 | 3.06 |
| Jumlah | 43.98 | 45.12 |

Perhitungan medoid cluster 1 iterasi 5 = D2<D9 maka medoid tetap. Kemudian bandingkan dengan selanjutnya yaitu D10 ditunjukkan pada tabel 4.13

Tabel 4.13 Cluster 1 iterasi 6

|  |  |  |
| --- | --- | --- |
|  | D2 | D10 |
| D1 | 5.28 | 3.05 |
| D2 | 0 | 2.28 |
| D3 | 5.06 | 5.61 |
| D4 | 5.02 | 3.4 |
| D5 | 7.52 | 8.63 |
| D6 | 1.55 | 3.80 |
| D7 | 9.88 | 11.0 |
| D8 | 5.34 | 4.24 |
| D9 | 2.05 | 3.06 |
| D10 | 2.28 | 0 |
| Jumlah | 43.98 | 45.07 |

Perhitungan medoid cluster 1 iterasi 5 = D2<D9 maka medoid tetap. Atau titik pusat pada cluster 1 yaitu D2. Sehingga mendapatkan hasil akhir cluster 1 sudah tidak berubah yang ditunjukkan pada tabel 4.14

Tabel 4.14 Hasil akhir cluster 1

|  |
| --- |
| Cluster 1 |
| Titik pusat D2 |
| D3 |
| D5 |
| D6 |
| D7 |
| D9 |
| D10 |

Setelah cluster 1 selesai dan jika medoid dari cluster 1 sudah tidak berubah maka selanjutnya mencari data yang tergabung dalam cluster 2 dengan cara yang sama pada tabel 4.15

Tabel 4.15 Hasil Clustering 2 Pertama

|  |
| --- |
| Cluster 2 |
| Clustering 2 (titik D4) |
| D1 |
| D8 |

Cluster 2 iterasi 1 : jumlah jarak data yang berada pada medoid ke-1 dan ke-2, jika jumlah jarak medoid ke-1 lebeih besar dari pada jumlah jarak medoid ke-2 maka medoid pada clster 1 menjadi data ke-2, namun jika sebaliknya jumlah pada medoid ke-1 lebih kecil dari pada jumlah ke-2 maka medoid tetap data ke-1 ditunjukkan pada tabel 4.16.

Tabel 4.16 Cluster 2 iterasi 1

|  |  |  |
| --- | --- | --- |
|  | D4 | D1 |
| D1 | 2.2 | 0 |
| D2 | 5.02 | 5.28 |
| D3 | 7.51 | 7.51 |
| D4 | 0 | 2.2 |
| D5 | 10.28 | 10.66 |
| D6 | 6.50 | 6.83 |
| D7 | 12.48 | 12.66 |
| D8 | 1.4 | 3.6 |
| D9 | 4.20 | 5.31 |
| D10 | 3.40 | 3.05 |
| Jumlah | 52.99 | 57.1 |

Perihtungan medoid cluster 2 iterasi 1 = D4<D1 maka medoid tetap. Kemudian bandingkan dengan selanjutnya yaitu D8 ditunjukkan pada tabel 4.17.

Tabel 4.17 Cluster 2 iterasi 2

|  |  |  |
| --- | --- | --- |
|  | D4 | D8 |
| D1 | 2.2 | 3.6 |
| D2 | 5.02 | 5.34 |
| D3 | 7.51 | 7.84 |
| D4 | 0 | 1.4 |
| D5 | 10.28 | 10.28 |
| D6 | 6.50 | 6.67 |
| D7 | 12.48 | 12.38 |
| D8 | 1.4 | 0 |
| D9 | 4.20 | 4.00 |
| D10 | 3.40 | 4.24 |
| Jumlah | 52.99 | 55.75 |

Perhitungan medoid cluster 2 iterasi 2 = D4<D8 maka medoid tetap . sehingga mendapatkan hasil akhir dari cluster 2 yang ditunjukan pada tabel 4.18.

Tabel 4.18 Hasil Akhir Cluster 2

|  |
| --- |
| Cluster 2 |
| Titik pusat D4 |
| D1 |
| D8 |

Dari perhitungan iterasi maka didapatakan hasil medoid setiap cluster, kemudian dilakukan perbandingan seluruh data dengan medoid yang dihasilkan yaitu D2 dan D4 seperti pada tabel 4.18.

Tabel 4.18 Cluster Kedua

|  |  |  |  |
| --- | --- | --- | --- |
|  | D2 | D4 | Cluster |
| D1 | 5.28 | 2.2 | 1 |
| D3 | 5.06 | 7.51 | 1 |
| D4 | 7.52 | 10.28 | 1 |
| D5 | 1.55 | 6.50 | 1 |
| D7 | 9.88 | 12.48 | 1 |
| D8 | 5.34 | 1.4 | 2 |
| D9 | 2.05 | 4.20 | 1 |
| D10 | 2.28 | 3.4 | 1 |

Kemudian dilakukan pengujian dengan Shilhoutte Coefficient seperti pada tabel 4.19

Tabel 4.19 Pengujian

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cluster |  | a(i) | b(i) | s(i) = 1-(a)/(b) |
| 1 | D3 | 1.27 | 2.32 | 0.46 |
|  | D5 | 1.70 | 3.13 | 0.46 |
|  | D6 | 1.81 | 0.35 | -4.17 |
|  | D7 | 2.45 | 5.72 | 0.58 |
|  | D9 | 0.33 | 1.67 | 0.81 |
|  | D10 | 1.61 | 0.44 | -2.65 |
| 2 | D1 | 0.64 | 1.58 | 0.6 |
|  | D8 | 0.64 | 1.85 | 0.66 |
| Hasil Shilhoutte Coefficient | | | | -3.25 |

* 1. **Implimentasi**

Implimentasi program “Penerapan Data Mining Untuk Clustering Penyakit Ispa Dengan Metode K-Medoids” ini menggunakan bahas pemograman python. Menghasilkan python dengan sebagai berikut:

* + 1. Selection

Tahap seleksi dilakukan pemilihan data yang berupa variabel-variabel yang dibutuhkan dari sekumpulan data penyakit ispa yaitu Berat badan, Kondisi lingkungan, Status gizi.

Sintak yang digunakan untuk selection data dapat dilihat pada tabel 4.1.

|  |  |
| --- | --- |
| 1 | d = drop[1:200][[4,5,6]] |
| 2 | d.columns = [‘Berat badan’,Kondisi lingkungan’,’Status gizi’] |
| 3 | D |

Listing 4.1 Selection data

Pada listing 4.1 baris 1 menjelaskan sintak untuk mengambil variabel 4,5,6 dari data 1 sampai 200. Baris 2 sintak untuk memanggil columns dari variabel yang sudah diseleksi yaitu berat badan, kondisi lingkungan, status gizi. Baris 3 sintak untuk menampung perintag sintak 1 dan 2.

* + 1. proses Analisis K-medoids

1. pada penelitian ini cluster yang ditentukan terdapat 2 cluster. 2

|  |  |
| --- | --- |
| 1 | range\_n\_clusters = [2] |
| 2 | semua\_hasil\_c=[] |

Listing 4.1 program untuk menentukan Cluster

baris 1 menjelaskan sintak untuk memasukkan jumlah cluster yang akan diproses, misal pada penelitian ini menggunakan 2 cluster. Baris 2 menjelaskan sintak untuk memanggil semua hasil cluster yang akan diproses.

1. Memilih titik pusat secara random sebanyak cluster yang ditentukan dimana titik ini yang akan menjadi pusat dari masing-masing cluster.

Listing program untuk menentukan dan menampilkan anggotanya dari cluster.

|  |  |
| --- | --- |
| 1 | for n\_clusters in range\_n\_clusters: |
| 2 | dist= pairwise\_distances(X, metric='euclidean') #membentuk euclidean |
| 3 | km=KMedoids(distance\_matrix=dist,n\_clusters=n\_clusters) |
| 4 | km.run(max\_iterations=8, tolerance=0.001) #mulai memproses k-medoids |
| 5 | hasilcluster = km.clusters |
| 6 | for key1, val1 in hasilcluster.items():  listval1 = list(val1) |
| 7 | for i in range(len(listval1)):  listcluster[listval1[i]] = a  satuitem={'Kluster':a, 'Berat badan':L[listval1[i]+1][0], 'Kondisi lingkungan':X[listval1[i]][0], 'Status gizi ':X[listval1[i]][1], } |
| 8  9 | listallcluster[a].append(satuitem)  a=a+1  for zz in range(n\_clusters):  hasil\_c['c'+str(zz)] = listallcluster[zz] |

Listing 4.3 menetukan centroid dan mencetak hasil anggota cluster.

menjelaskan sintak untuk menetukan centroid dan mencetak hasil anggota cluster. Pada baris 2 membentuk centroid dengan menggunakan Euclidean tertentu. Baris 4 sintak untuk mulai memproses k-medoids. Baris 5 sintak untuk menampung semua sintak 1,2,3,4. Baris 6 sintak untuk memproses semua item yang ada di hasil k-medoids. Baris 7 sintak untuk memproses item yang ada di suatu cluster, misalnya cluster 0. Baris 8 sintak untuk memasukan hasil mudah dibaca kedalam list. Baris 9 sintak untuk memasukan hasil list yang mudah dibaca kedalam bentuk dictionary.

1. Pengujian shillhoutte Coeficient

|  |  |
| --- | --- |
| 1 | silhouette\_avg = silhouette\_score(X, listcluster) |

Listing 4.4 Menghitung sillhoutte

baris 1 sintak untuk proses menghitung sillhoutte dari hasil k-medoids diatasnya.

* 1. **Pengujian Akurasi**

Pengujian akurasi menggunakan Shilhoutte Coeficient dilakukan untuk mengetahui kelompok yang baik untuk diterapkan dengan jumlah data yang ada yaitu 200 berdasarkan hasil dari proses pengelompokkan dengan metode K-medoids. Jika hasil Shilhoutte Coeficient semakin mendekati 1, maka semakin baik kualitas kelompoknya. Sedangkan, jika hasil Shilhoutte Coeficient semakin mendekati -1, maka semakin kurang baik kualitas kelompoknya. Hasil dari Shilhoutte Coeficient dari data 200, pengujian sejumlah 5 cluster dilakukan di program agar dapat dibandingkan dengan pengelompokkan menggunakan metode K-Medoids dapat dilihat pada tabel 4.24 Hasil shilhoutte Coeficient.

Tabel 4.20 Hasil shilhoutte Coeficient.

|  |  |
| --- | --- |
| Cluster | Hasil Pengujian (Si) |
| 2 | 0.344 |
| 3 | 0.617 |
| 4 | 0.245 |
| 5 | 0.193 |
| 6 | 0.456 |

Dari perhitungan di atas didapat hasil Shilhouette Coeficient adalah terbesar

0.617 hasil tersebut meunjukan pengelompokan paling baik adalah apabila menggunakan 3 cluster.

* 1. Evaluasi pola dan Representasi Pengetahuan

Proses clustering mengelompokan data penyakit ispa ke suatu kelompok, berdasarkan berat badan, kondisi lingkungan, status gizi. Proses iterasi clustering data penyakit ispa sebanyak 6 kali, pada iterasi 1,2,3,4,5,6 tidak ada perubahan data karena hasil yang dicapai sudah konvergen. Hasil data penyakit ispa yang sudah di clustering menjadi 2 yaitu penyakit ispa berdasarkan status gizi redah dan baik. Adapun data penyakit ispa yang terbagi menjadi 2 cluster adalah sebagai berikut:

1. Berdasarlan hasil kluster dapat dianalisis kelompok 2 kategori kurang baik karena di lihat dari titik pusat memiliki total lebih tinggi yaitu 72 sedangkan kelompok 1 yaitu 57. dimana kelompok 1 termasuk kelompok yang memiliki kriteria kurang baik dalam semua segi penilaian yaitu berat badan yang tidak menentu, kondisi lingkungan yang tidak menentu , status gizi rendah. sedangkan kelompok 2 termasuk kriteria baik dilihat berat badan tidak menentu, kondisi lingkungan baik, status gizi tinggi.
2. Pola kelompok tersebut dilihat dari hasil shilhoutte Coeficient adalah sebesar 0.617. Hasil tersebut tergolong rendah karena shilhoutte Coeficient memiliki range nilai antara 0-1. Semakin mendekati 1 hasil kelompoknya semakin baik.
3. Hasil Shilhoutte Coeficient yang rendah disebabkan oleh data yang ada terlalu spesifik mengenai Ispa sehingga terjadi ketidak konsistenan data.
4. Lalu dari pemilihan metode juga berpengaruh terhadap rendahnya nilai Shilhoutte Coeficient, karena K-Medoids merupakan pengelompokkan dengan pendekatan partisi (patitional clustering) yang bekerja efektif untuk dataset yng kecil, tetapi tidak berjalan baik untuk dataset yang besar. Pengelompokkan menggunakan metode K-Medoids tidak dipengaruhi oleh data yang ekstrem, akan tetapi dengan jumlah variable yang besar dan memiliki indeks nilai yang berbeda sehingga kurang optimal diterapkan. Pengelompokkan Ispa lebih cocok menggunakan metode yang efektif dalam mengelompokkan dataset dengan jumlah besar, serta dapat menggunakan metode yang dapat memecah dataset dengan jumlah besar menjadi beberapa sample yang mirip dengan data asli.
5. Metode pengelompokkan K-Medoids menggunakan Euclidean Distance di penelitian ini menghasilkan kelompok yang kurang bagus dan kurang cocok untuk pengelompokkan data yang memiliki index nilai yang berbeda-beda.