



PENERAPAN DATA MINING K-MEANS CLUSTERING UNTUK MENGELOMPOKKAN BERBAGAI JENIS MERK LAPTOP

Putri Ulil Fatma Aulia¹⁾, Sudin Saepudin²⁾

^{1, 2)}Program Studi Sistem Informasi, Universitas Nusa Putra

Jl. Raya Cibolang No. 21 Cibolang Kaler, Cisaat, Sukabumi, Jawa Barat 43152

e-mail: putri.ulil_si18@nusaputra.ac.id¹⁾, sudinsaepudin@nusaputra.ac.id²⁾

* Korespondensi: e-mail: sudinsaepudin@nusaputra.ac.id

ABSTRAK

Laptop sudah menjadi kebutuhan yang tak terpisahkan dari masyarakat di zaman sekarang ini. Laptop adalah komputer bergerak (bisa dipindahkan dengan mudah) yang berukuran relatif kecil dan ringan, beratnya berkisar dari 1-6 kg, tergantung ukuran, bahan, dari spesifikasi laptop tersebut. Dengan berkembangnya teknologi mobile saat ini menjadikan jenis dan merk laptop semakin banyak. Banyaknya pilihan laptop tidak otomatis memudahkan kita untuk membeli yang terbaik karena banyak masyarakat yang kurang tahu dari spesifikasi laptop. Tujuan penelitian ini menerapkan metode k-means Clustering untuk mengelompokkan berbagai merk jenis laptop sehingga dapat mempermudah konsumen untuk membeli dengan kualitas yang terbaik berdasarkan pengujian yang dilakukan mampu memberikan hasil yang baik sesuai dengan perhitungan yang digunakan dan dapat membantu konsumen yang akan membeli. Kesimpulan yang diperoleh dari penelitian ini adalah hubungan positif yang sangat signifikan antara persepsi konsumen terhadap merek, kualitas dengan minat membeli.

Kata Kunci: k-means, Clustering, Laptop

ABSTRACT

Laptops have become an inseparable need in today's society. Laptops are mobile computers (can be moved easily) that are relatively small and light in weight, weighing in the range of 1-6 kg, depending on the size, material, and specifications of the laptop. With the development of mobile technology today, there are more and more types and brands of laptops. The large selection of laptops does not automatically make it easier for us to buy the best, because many people do not know about laptop specifications. The purpose of this study is to apply the k-means clustering method to group various brands of laptop types so that it can make it easier for consumers to buy the best quality. The conclusion obtained from this study is a very significant positive relationship between consumer perceptions of brands, quality and buying interest.

Keywords: k-means, Clustering, Laptops

I. PENDAHULUAN

Kebutuhan teknologi saat ini meningkat seiring dengan berkembangnya teknologi. Perkembangan teknologi saat ini menjadikan pengguna laptop sudah menjadi kebutuhan mendasar yang harus dimiliki untuk membantu dalam menyelesaikan suatu pekerjaan sebagai pengganti personal computer. Hal tersebut mengakibatkan semakin banyak bermunculan berbagai jenis merk laptop yang saling bersaing untuk dapat memenuhi kebutuhan konsumen saat ini dengan berbagai jenis spesifikasi yang dimilikinya. Namun tidak setiap konsumen memahami tentang spesifikasi dari laptop yang disesuaikan dengan kebutuhan dan anggaran yang ditentukan. Oleh karena itu diperlukan suatu metode yang dapat membantu konsumen dalam menentukan laptop yang akan dibeli sesuai dengan kebutuhan dan anggaran. Penelitian ini mengusulkan sebuah metode K-Means Clustering yang bertujuan untuk mengelompokkan data yang ada kedalam berbagai kelompok, dimana data dalam suatu kelompok mempunyai karakteristik yang sama satu sama lainnya.

II. TINJAUAN PUSTAKA

A. Data Mining

Data mining merupakan proses menemukan korelasi baru yang bermanfaat, pola dan trend dengan menambang sejumlah repositori data dalam jumlah besar. Data mining semakin menyebar dan berkembang dengan pesat belakangan ini karena kemampuannya dalam menambang pola bermanfaat dan trend dari basis data yang sudah ada [1]. Tujuan utama data mining adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki [3].

B. Clustering

Salah satu pengelompokan dalam data mining adalah metode clustering. Clustering merupakan salah satu metode data mining yang bersifat tanpa arahan yang memiliki arti bahwa karakteristik tiap kluster tidak ditentukan sebelumnya melainkan berdasarkan kemiripan atribut atribut dari suatu kelompok atau cluster [2]. dengan menggunakan clustering maka dapat mengklasifikasikan daerah yang padat, menemukan pola-pola distribusi secara keseluruhan, dan menemukan keterkaitan yang menarik antara atribut data. pada dasarnya Clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik antara data satu dengan data yang lain [4].

C. Algoritma K-Means

Algoritma k-means merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah k cluster yang sudah ditetapkan di awal. tujuan pengelompokan ini adalah untuk meminimalkan fungsi objek yang di set dalam proses pengelompokan [5] Algoritma k-means sangat terkenal karena kemudahan dan kemampuannya untuk mengkluster data yang besar dan data outlier dengan sangat cepat. Dalam algoritma k-means, setiap data harus termasuk ke cluster tertentu dan bisa dimungkinkan bagi setiap data yang termasuk cluster tertentu pada satu tahapan proses, pada tahapan berikutnya berpindah ke cluster lainnya. Posisi pusat cluster akan dihitung kembali sampai semua komponen data digolongkan ke dalam tiap pusat cluster dan terakhir akan terbentuk posisi pusat cluster yang baru. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori Euclidean yang dirumuskan sebagai berikut :

$$D(i,j) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \dots + (Xki - Xkj)^2} \dots (1)$$

Dimana :

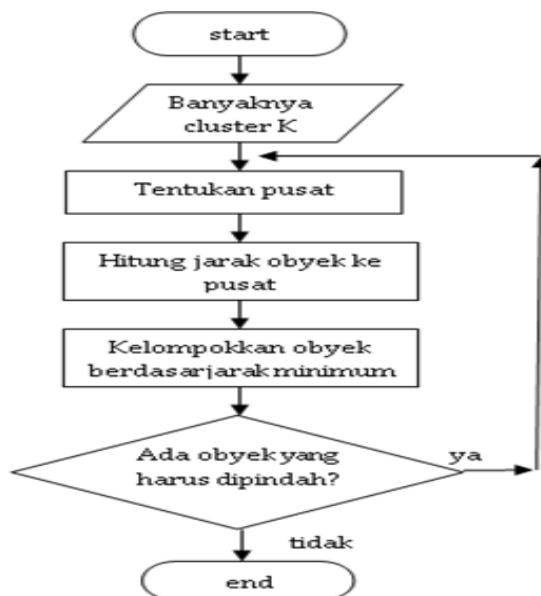
$D(i,j)$ = jarak data ke i ke pusat cluster j

Xki = data ke i pada atribut data ke k

Xkj = titik pusat ke j pada atribut ke k

III. METODOLOGI PENELITIAN

Penelitian ini dilakukan dalam beberapa tahap yaitu :



Gambar 1. Flowchart algoritma K-Means

Gambar 1. diatas menunjukkan langkah – langkah menentukan clustering dengan menggunakan metode k-means. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data ke tiap pusat cluster. Untuk menghitung jarak jarak semua data ke setiap titik pusat cluster dapat menggunakan jarak Euclidean yang dirumuskan sebagai berikut :

1. Menentukan nilai k sebagai jumlah Cluster yang akan dibentuk.
2. Inisialisasi k centroid (titik pusat Cluster) awal secara random.



3. Menghitung jarak setiap data ke masing-masing centroid menggunakan rumus korelasi antara dua objek.
4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
5. Apabila masih ada data yang berpindah cluster atau perubahan nilai centroid.

$$d = \sqrt{(X1 - X2)^2 + (y1 - y2)^2} \dots\dots\dots(2)$$

Keterangan

- d = Jarak
X1 = Koordinat latitude 1
X2 = Koordinat latitude 2
Y1 = Koordinat Longitude 1
Y2 = Koordinat Longitude 2

Koordinat latitude dan longitude dapat dijadikan variabel guna melakukan perhitungan jarak antara dua buah titik apabila dibentangkan garis lurus diantara keduanya.

IV. HASIL DAN PEMBAHASAN

A. Menentukan nilai k sebagai jumlah cluster yang akan dibentuk

Data yang dianalisis yaitu data dari 25 merk Laptop.

Tabel.1. Data Laptop

NO	Nama	Parameter	
		RAM	HARDDISK
1	Lenovo ideapad 330	4	500
2	Lenovo ideapad 320	4	1000
3	Lenovo V14	4	1000
4	Asus VivoBook Flip 12	4	1000
5	Asus E402WA	4	500
6	Acer Aspire 3	4	500
7	Lenovo Ideapad 110	4	500
8	Dell inspiron 11	4	500
9	Dell Latitude 3330	4	320
10	Asus aspire D270	2	320
11	Asus X2	2	320
12	Acer Aspire one 722	2	320
13	Asus A43E	2	500
14	Asus Aspireone Happy	2	500
15	Toshiba satellite Lg45	2	320
16	Hp Mini 110	2	320
17	Asus VivoBook X540YA	2	500
18	Dell Latitude 3410	8	1000
19	Hp 14S	8	512
20	Acer Aspire Travelmate	8	500
21	Lenovo ideapad S340	8	512
22	Lenovo V11	8	1000
23	Asus R410 MA-BV208t	8	1000
24	Dell Inspiron 15-3567	8	1000
25	Hp NoteBook 15	8	1000

B. Inisialisasi k centroid (titik pusat Cluster) awal secara random.

Berdasarkan data dari tabel diatas, maka banyaknya cluster yang dibentuk tiga ($k=3$). Banyaknya cluster harus lebih kecil dari pada banyaknya data ($k < n$).

Tabel.2. Titik Pusat Awal Setiap Cluster

Jumlah Cluster	3		
Pusat Cluster	(random)		
Cluster	Nama Laptop	RAM	HARDISK
C1	Lenovo Ideapad 330	4	500
C2	Toshiba Satellite	2	320
C3	Lenovo V11	8	1000

Inisialisasi centroid dataset pada tabel dataset diatas adalah $C1 = \{ 4, 500 \}$, $C2 = \{ 2, 320 \}$, $C3 = \{ 8, 1000 \}$. Inisialisasi centroid ini ditentukan secara random.

Bentuk pengulangan berikutnya, centroid baru dihitung dengan menghitung nilai rata-rata data pada setiap cluster. Jika centroid yang baru berbeda dengan centroid sebelumnya maka dilanjutkan ke langkah berikutnya, begitupun sebaliknya jika centroid baru sama dengan yang sebelumnya maka clustering selesai.

C. Menghitung jarak setiap data ke masing-masing centroid menggunakan rumus korelasi antara dua objek.

Tempatkan setiap data pada cluster, sehingga data akan dimasukkan kedalam cluster yang memiliki jarak paling dekat dengan titik pusat dari setiap cluster, untuk mengetahui cluster mana yang paling dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap cluster. berikut perhitungan jarak dari data laptop pertama ke pusat cluster pertama :

- Jarak data laptop dengan cluster 1 (4 ,500)

$$\begin{aligned}
 1. D(x_1, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4)^2 + (500 - 500)^2} = 0 \\
 2. D(x_2, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4)^2 + (1000 - 500)^2} = 500 \\
 3. D(x_3, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4)^2 + (1000 - 500)^2} = 500 \\
 4. D(x_4, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4)^2 + (1000 - 500)^2} = 500 \\
 5. D(x_5, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4)^2 + (500 - 500)^2} = 0
 \end{aligned}$$

Lakukan perhitungan yang sama sampai data ke 25

- Jarak data laptop dengan cluster 2 (2, 320)

$$\begin{aligned}
 1. D(x_1, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2)^2 + (500 - 320)^2} = 180,0111 \\
 2. D(x_2, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2)^2 + (1000 - 320)^2} = 680,0029 \\
 3. D(x_3, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2)^2 + (1000 - 320)^2} = 680,0029 \\
 4. D(x_4, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2)^2 + (1000 - 320)^2} = 680,0029 \\
 5. D(x_5, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2)^2 + (500 - 320)^2} = 180,0111
 \end{aligned}$$

Lakukan perhitungan yang sama sampai data ke 25

- Jarak data laptop dengan cluster 3 (8, 1000)

$$\begin{aligned}
 1. D(x_1, c_3) &= \sqrt{(a_1 - c_3a)^2 + (b_1 - c_3b)^2} = \sqrt{(4 - 8)^2 + (500 - 1000)^2} = 500,0159 \\
 2. D(x_2, c_3) &= \sqrt{(a_1 - c_3a)^2 + (b_1 - c_3b)^2} = \sqrt{(4 - 8)^2 + (1000 - 1000)^2} = 4 \\
 3. D(x_3, c_3) &= \sqrt{(a_1 - c_3a)^2 + (b_1 - c_3b)^2} = \sqrt{(4 - 8)^2 + (1000 - 1000)^2} = 4 \\
 4. D(x_4, c_3) &= \sqrt{(a_1 - c_3a)^2 + (b_1 - c_3b)^2} = \sqrt{(4 - 8)^2 + (1000 - 1000)^2} = 4 \\
 5. D(x_5, c_3) &= \sqrt{(a_1 - c_3a)^2 + (b_1 - c_3b)^2} = \sqrt{(4 - 8)^2 + (500 - 1000)^2} = 500,0159
 \end{aligned}$$



- Untuk seterusnya hitung jarak pada setiap baris data, dan hasilnya seperti tabel dibawah :

Tabel.3. Data table

Jarak ke pusat Cluster						
Pusat cluster	Nama	Parameter		C1	C2	C3
		RAM	HARDDISK	Dc1	Dc2	Dc3
1	Lenovo ideapad 330	4	500	0	180,0111108	500,0159997
2	Lenovo ideapad 320	4	1000	500	680,0029412	4
3	Lenovo V14	4	1000	500	680,0029412	4
4	Asus VivoBook Flip 12	4	1000	500	680,0029412	4
5	Asus E402WA	4	500	0	180,0111108	500,0159997
6	Acer Aspire 3	4	500	0	180,0111108	500,0159997
7	Lenovo Ideapad 110	4	500	0	180,0111108	500,0159997
8	Dell inspiron 11	4	500	0	180,0111108	500,0159997
9	Dell Latitude 3330	4	320	180	2	680,0117646
10	Asus aspire D270	2	320	180,0111108	0	680,0264701
11	Asus X2	2	320	180,0111108	0	680,0264701
12	Acer Aspire one 722	2	320	180,0111108	0	680,0264701
13	Asus A43E	2	500	2	180	500,0359987
14	Asus Aspire One Happy	2	500	2	180	500,0359987
15	Toshiba satellite Lg45	2	320	180,0111108	0	680,0264701
16	Hp Mini 110	2	320	180,0111108	0	680,0264701
17	Asus VivoBook X540YA	2	500	2	180	500,0359987
18	Dell Latitude 3410	8	1000	500,0159997	680,0264701	0
19	Hp 14S	8	512	12,64911064	192,0937271	488
20	Acer Aspire Travelmate	8	500	4	180,0999722	500
21	Lenovo ideapad S340	8	512	12,64911064	192,0937271	488
22	Lenovo V11	8	1000	500,0159997	680,0264701	0
23	Asus R410 MA-BV208t	8	1000	500,0159997	680,0264701	0
24	Dell Inspiron 15-3567	8	1000	500,0159997	680,0264701	0
25	Hp NoteBook 15	8	1000	500,0159997	680,0264701	0

Tabel.4. Data Tabel Pusat Cluster Baru

C1		C2		C3	
RAM	HARDDISK	RAM	HARDDISK	RAM	HARDDISK
4	500				
				4	1000
				4	1000
				4	1000
4	500				
4	500				
4	500				
4	500				
		4	320		
		2	320		
		2	320		
		2	320		
2	500				



2	500				
		2	320		
		2	320		
2	500				
				8	1000
8	512				
8	500				
8	512				
				8	1000
				8	1000
				8	1000
				8	1000
4,5455	502,18182	2,333333333	320	6,5	1000

Kelompokkan data sesuai dengan clusternya yaitu data yang memiliki jarak terpendek. Misalnya: karena $d_3 < d_2 < d_1$ maka x_1 masuk kedalam cluster 3. Pada tabel diatas data $N = 1$ masuk kedalam cluster 3 karena $d_3 < d_2 < d_1$, sedangkan $N = 14, 15, 16$ Sampai dengan 22 masuk kedalam cluster 2.

Setelah mendapatkan label cluster untuk masing masing data $n=1, 2, 3, 4, 5 \dots 25$ maka dicari nilai rata ratanya dengan menjumlahkan seluruh anggota masing masing cluster dan dibagi jumlah anggotanya.

Tabel.5. Data table Nilai

Cluster	RAM	HARDISK
C1	4,5454545	502,181818
C2	2,3333333	320
C3	6,5	1000

Nilai Rata rata Centroid pada pengulangan ke 1

Dari hasil literasi 1 dapat dilihat centroid baru yang dihasilkan dari iterasi 1 dengan titik pusat awal tidak sama, maka perlu dilakukan iterasi kedua.

Perhitungan jarak euclidean pada iterasi 2

- Jarak data laptop dengan cluster 1 (4,5454 ,502,1818)

$$\begin{aligned}
 1. D(x_1, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4,5454)^2 + (500 - 502,1818)^2} = 2,2489 \\
 2. D(x_2, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4,5454)^2 + (1000 - 502,1818)^2} = 497,8184 \\
 3. D(x_3, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4,5454)^2 + (1000 - 502,1818)^2} = 497,8184 \\
 4. D(x_4, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4,5454)^2 + (1000 - 502,1818)^2} = 497,8184 \\
 5. D(x_5, c_1) &= \sqrt{(a_1 - c_1a)^2 + (b_1 - c_1b)^2} = \sqrt{(4 - 4,5454)^2 + (500 - 502,1818)^2} = 2,2489
 \end{aligned}$$

Lakukan perhitungan yang sama sampai data ke 25

- Jarak data laptop dengan cluster 2 (2,3333, 320)

$$\begin{aligned}
 1. D(x_1, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2,3333)^2 + (500 - 320)^2} = 180,0077 \\
 2. D(x_2, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2,3333)^2 + (1000 - 320)^2} = 680,0024 \\
 3. D(x_3, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2,3333)^2 + (1000 - 320)^2} = 680,0024 \\
 4. D(x_4, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2,3333)^2 + (1000 - 320)^2} = 680,0024 \\
 5. D(x_5, c_2) &= \sqrt{(a_1 - c_2a)^2 + (b_1 - c_2b)^2} = \sqrt{(4 - 2,3333)^2 + (500 - 320)^2} = 180,0077
 \end{aligned}$$

Lakukan perhitungan yang sama sampai data ke 25



- Jarak data laptop dengan cluster 3 (6,5 , 1000)

- $D(x1,c3) = \sqrt{(a1 - c3a)^2 + (b1 - c3b)^2} = \sqrt{(4 - 6,5)^2 + (500 - 1000)^2} = 500,0062$
- $D(x2,c3) = \sqrt{(a1 - c3a)^2 + (b1 - c3b)^2} = \sqrt{(4 - 6,5)^2 + (1000 - 1000)^2} = 2,5$
- $D(x3,c3) = \sqrt{(a1 - c3a)^2 + (b1 - c3b)^2} = \sqrt{(4 - 6,5)^2 + (1000 - 1000)^2} = 2,5$
- $D(x4,c3) = \sqrt{(a1 - c3a)^2 + (b1 - c3b)^2} = \sqrt{(4 - 6,5)^2 + (1000 - 1000)^2} = 2,5$
- $D(x5,c3) = \sqrt{(a1 - c3a)^2 + (b1 - c3b)^2} = \sqrt{(4 - 6,5)^2 + (500 - 1000)^2} = 500,0062$

Tabel.6. Data table Pusat cluster baru

C1		C2		C3	
RAM	HARDDISK	RAM	HARDDISK	RAM	HARDDISK
4	500				
				4	1000
				4	1000
				4	1000
4	500				
4	500				
4	500				
4	500				
		4	320		
		2	320		
		2	320		
		2	320		
2	500				
2	500				
		2	320		
		2	320		
2	500				
				8	1000
8	512				
8	500				
8	512				
				8	1000
				8	1000
				8	1000
				8	1000

Setelah mendapatkan label cluster untuk masing masing data n=1,2,3,4,5...29 maka dicari nilai ratanya dengan menjumlahkan masing masing cluster dan dibagi jumlah anggotanya .

Tabel.7. Data table nilai

Cluster	RAM	HARDDISK
C1	4,5454545	502,181818
C2	2,3333333	320
C3	6,5	1000

Nilai rata rata centroid pada pengulangan ke 2

Karena Centroid tidak mengalami perubahan (sama dengan centroid sebelumnya) maka proses clustering selesai.



Hasil :

Table .8 Data table cluster

N0	Nama	Kelompok Cluster
1	Lenovo ideapad 330	C1
2	Lenovo ideapad 320	C3
3	Lenovo V14	C3
4	Asus VivoBook Flip 12	C3
5	Asus E402WA	C1
6	Acer Aspire 3	C1
7	Lenovo Ideapad 110	C1
8	Dell inspiron 11	C1
9	Dell Latitude 3330	C2
10	Asus aspire D270	C2
11	Asus X2	C2
12	Acer Aspire one 722	C2
13	Asus A43E	C1
14	Asus Aspireone Happy	C1
15	Toshiba satellite Lg45	C2
16	Hp Mini 110	C2
17	Asus VivoBook X540YA	C1
18	Dell Latitude 3410	C3
19	Hp 14S	C1
20	Acer Aspire Travelmate	C1
21	Lenovo ideapad S340	C1
22	Lenovo V11	C3
23	Asus R410 MA-BV208t	C3
24	Dell Inspiron 15-3567	C3
25	Hp NoteBook 15	C3

V. KESIMPULAN

Berdasarkan penelitian ini dapat disimpulkan sebagai berikut :

1. Telah berhasil dilakukan pengelompokan data laptop menggunakan algoritma k means clustering menjadi 3 kelompok yaitu untuk kelompok 1 berjumlah 11 data, kelompok 2 berjumlah 6 data, kelompok 3 berjumlah 8 data.
2. Penentuan titik pusat awal (centroid) sangat berpengaruh terhadap jumlah iterasi yang akan dihitung.
3. Berdasarkan penelitian ini dapat dijadikan acuan yang dapat dilakukan konsumen untuk memilih laptop sesuai anggaran dan kebutuhan.

DAFTAR PUSTAKA

- [1] Fina Nasari, Charles Jhony Manto Sianturi “Penerapan Algoritma K means Clustering untuk Pengelompokan Penyebaran diare di kabupaten Langkat” Cotigo Smart Journal, Vol. 2/No.2/desember 2016.
- [2] Roni Setiawan “ Penerapan Data Mining Menggunakan Algoritma k means Clustering untuk Menentukan Strategi Promosi Mahasiswa Baru “ Jurnal Lentera ICT , Vol.3 No.1, Mei 2016.



- [3] susanto, Santi dan dedi Suryadi. 2010. Pengantar Data Mining Menggali Pengetahuan dari Bongkahan Data. Yogyakarta: Andi.
- [4] Lilis Suryani “ Pengelompokkan data kriminal pada poldasu menentukan pola rawan tindak kriminal menggunakan data mining algoritma k means clustering”, Vol. 1/No.2/ Januari 2020.
- [5] fauziah Nur “Penerapan Algoritma k means pada siswa baru sekolah menengah kejuruan untuk clustering jurusan” Jurnal Nasional Informatika Dan Teknologi Jaringan, Vol.1, No.2 , Maret 2017.