

Penerapan Data Mining Untuk Pengelompokan Minat Konsumen Terhadap Pengguna Jasa Pengiriman Pada PT. Jalur Nugraha Ekakurir (JNE) Pematangsiantar

Rudi Wardana Nasution^{1*}, Suhada², Ika Okta Kirana³, Indra Gunawan⁴, Ika Purnama Sari⁵

^{1,3,4,5} STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

² AMIK Tunas Bangsa, Pematangsiantar, Indonesia

Email: ^{1*} rudiwardananst@gmail.com, ² suhada.atb@gmail.com, ³ ikaoktakirana123@gmail.com, ⁴ indra@amiktunasbangs.ac.id,

⁵ ikapurnamasari1319@gmail.com

Abstrak—*Datamining* merupakan suatu metode untuk menemukan pengetahuan dalam suatu tumpukan data yang cukup besar dengan proses menggali dan menganalisa sejumlah data yang sangat besar dalam memperoleh sesuatu yang benar, baru dan bermanfaat sehingga dapat ditemukan suatu corak atau pola dalam data tersebut Pelayanan (*customer service*) secara umum adalah setiap kegiatan yang diperuntukkan atau ditujukan untuk memberikan kepuasan kepada pelanggan, melalui pelayanan ini keinginan dan kebutuhan pelanggan dapat terpenuhi. Dalam Kamus Bahasa Indonesia dijelaskan bahwa pelayanan adalah sebagai usaha melayani kebutuhan orang lain, sedangkan melayani yaitu membantu menyiapkan (membantu apa yang diperlukan seseorang). pada hakekatnya pelayanan adalah serangkaian kegiatan yang merupakan proses. Sebagai proses pelayanan berlangsung secara rutin dan berkesinambungan meliputi seluruh kehidupan orang dalam masyarakat, proses pemenuhan kebutuhan melalui aktivitas orang lain termasuk juga JNE merupakan perusahaan yang menawarkan pelayanan yang bergerak dalam bidang pengiriman dan logistik yang bermarkas di Jakarta, Indonesia. Nama resminya adalah Tiki Jalur Nugraha Ekakurir (Tiki JNE) *AlgoritmaK-Means* merupakan *Algoritma* pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah *K cluster* yang sudah ditetapkan di awal.

Kata Kunci: *Data Mining*, *Customer Service*, *Algoritma K-Means*

Abstract—*Datamining* is a method for finding knowledge in a large enough data pile by the process of digging and analyzing a very large amount of data in obtaining something that is true, new and useful so that it can be found in a pattern or pattern in the data. *Customer service* in general is every activity intended or intended to give satisfaction to customers, through this service the wishes and needs of customers can be met. In the Indonesian Dictionary explained that service is an attempt to serve the needs of others, while serving is to help prepare (help what someone needs). essentially service is a series of activities that constitute a process. As the service process takes place routinely and continuously covers the entire life of people in the community, the process of meeting needs through the activities of others including JNE is also a company that offers services engaged in shipping and logistics based in Jakarta, Indonesia. The official name is Tiki Path Nugraha Ekakurir (Tiki JNE) *K-Means Algorithm* is an iterative clustering algorithm that partitions data sets into a number of *K clusters* that have been set at the beginning.

Keywords: *Data Mining*, *Customer Service*, *K-Means Algorithm*

1. PENDAHULUAN

Datamining merupakan suatu metode untuk menemukan pengetahuan dalam suatu tumpukan data yang cukup besar dengan proses menggali dan menganalisa sejumlah data yang sangat besar dalam memperoleh sesuatu yang benar, baru dan bermanfaat sehingga dapat ditemukan suatu corak atau pola dalam data tersebut [1]. *Datamining* adalah salah satu cabang ilmu kecerdasan buatan dalam penggalian pola-pola untuk mengubah data menjadi informasi. *Datamining* merupakan proses yang mempekerjakan satu atau lebih teknik pembelajaran *computer* untuk menganalisa dan mengekstraksi pengetahuan *knowledge* secara otomatis [2]. Istilah *datamining* mulai dikenal sejak tahun 1990, munculnya *datamining* didasarkan pada jumlah data yang tersimpan dalam basis data yang semakin besar [3].

Dalam hal ini pihak Kantor JNE Kota Pematangsiantar penulis berencana mengelompokkan kembali Data minat konsumen terhadap jasa pengiriman yang ada pada kantor JNE agar dapat menarik kembali minat masyarakat untuk Jasa pengiriman pada JNE Tersebut. Untuk mempermudah dalam pengelompokan minat konsumen terhadap jasa pengiriman maka penelitian ini menggunakan metode *datamining*. Salah satu tugas utama *datamining* adalah *Clustering*. *Clustering* adalah proses pengorganisasian objek ke dalam class/kelompok dengan cara mencari kemiripan setiap objek [4].

Berdasarkan latar belakang masalah di atas, penulis menerapkan metode data mining dalam pengelompokan minat konsumen terhadap pengguna jasa pengiriman pada PT. Jalur Nugraha Ekakurir (JNE) Pematangsiantar. Diharapkan penelitian ini dapat membantu pihak kantor JNE Pematangsiantar dalam mengelompokkan minat konsumen terhadap jasa pengiriman sehingga dapat menarik kembali minat masyarakat untuk menggunakan jasa pengiriman pada kantor JNE Pematangsiantar sehingga dapat lebih meningkatkan keinginan masyarakat untuk selalu menggunakan jasa tersebut. Penelitian ini juga diharapkan dapat menjadi referensi bagi peneliti selanjutnya yang berkaitan dengan pengguna *Algoritma* yang digunakan.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Datamining merupakan salah satu metode dalam sistem cerdas yang mengolah data dalam jumlah besar. *Datamining* umumnya digunakan dalam sebuah proses analisa data untuk mengetahui suatu pola dari kelompok data yang tersembunyi[5]. Berikut beberapa penjelasan mengenai *datamining*. *Datamining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machinelearning* untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar[6].

Datamining dapat juga didefinisikan sebagai gabungan dari suatu pengalaman, nilai, informasi kontekstual dan juga pandangan pakar yang memberikan suatu *framework* untuk mengevaluasi dan menciptakan pengalaman baru dan informasi[7]. *Datamining* adalah bagian dari tahapan proses KDD. *Data Mining* atau *Knowledge Discovery in Database* (KDD) adalah penyaringan data secara implisit dimana sebelumnya tidak diketahui terdapatnya informasi yang potensial [8].

2.2 Algoritma K-Means

K-Means adalah suatu metode penganalisaan data atau metode *Datamining* yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi [9]. Berikut Alur penyelesaian dari metode *K-Means*[3]:

- Tentukan berapa jumlah kluster yang ingin ditetapkan pusat cluster k.
- Menggunakan jarak euclidean dan kemudian menghitung setiap data ke pusat cluster.

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (1)$$

- Kategorikan data ke dalam cluster dengan jarak yang terpendek dengan menggunakan persamaan

$$\text{Min } \sum_k^k = d_{ik} = d_{ik} \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (2)$$

- Menghitung pusat cluster dengan menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (3)$$

Dengan : $x_{ij} \in$ Kluster ke - k p = banyak member kluster ke - k

- Silahkan ulangi langkah dua sampai empat sehingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

3. HASIL DAN PEMBAHASAN

Data yang digunakan adalah data tahun 2019 dengan kurun waktu 3 bulan yaitu September, Oktober, dan November dengan cabang sebanyak 27 kota/kab/kec. Pengumpulan data dalam penelitian ini dilakukan dengan mengambil data yang telah ada di PT.JNE dan melakukan wawancara dengan pihak PT. JNE Pematangsiantar

Tabel 1. Data Penelitian
(Sumber : PT.JNE Pematangsiantar)

No	Kota/Kab/Kec	September	Oktober	November
1	Siantar Martoba	1900	2047	2500
2	Siantar Barat	1805	1306	1385
3	Siantar Marihat	1025	2050	2113
4	Siantar Utara	544	449	301
5	Siantar Selatan	702	706	760
6	Siantar Marimbun	1051	931	1139
7	Siantar Sitalasari	800	794	1102
8	Raya	649	599	824
9	Bandar	691	685	720
10	Dolok Batunanggar	1357	1051	1244
11	Girsang Sipangan Bolon	221	183	266
12	Pematang Bandar	589	502	504
13	Tanah Jawa	1996	1561	1772
14	Tapian Dolok	1424	1323	1632
15	Bandar Huluan	872	939	885
16	Bandar Masilam	1330	1277	1405
17	Bosar Maligas	87	144	191
18	Gunung Malela	783	774	761
19	Gunung Maligas	492	419	421
20	Hutabayu Raja	1445	1095	1015
21	Jorlang Hataran	374	478	206
22	Pane	266	655	596
23	Raya Kahean	1114	1106	1500
24	Sidamanik	917	871	891
25	Silimakuta	898	864	847
26	Ujung Padang	1119	960	1074

No	Kota/Kab/Kec	September	Oktober	November
27	Serbelawan	1857	1586	2149

3.1 Perhitungan K-Means

Untuk mempermudah dalam proses pengolahan maka data yang digunakan dalam proses k-means adalah total penjualan dari bulan september-november. Selanjutnya masuk dalam tahap perhitungan menggunakan metode *k-means*:

1. Penentuan Pusat Cluster Awal

Penentuan titik *cluster* ini dilakukan dengan mengambil nilai terbesar (C1), nilai terkecil untuk (C2) sebagai berikut:

Tabel 2. Pusat Cluster Awal

Cluster	September	Oktober	November
C1	1996	2050	2500
C2	87	144	191

2. Perhitungan Jarak Cluster

Untuk menghitung jarak antara data dengan pusat cluster memakai persamaan berikut:

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

Cluster Tinggi:

$$C_{(SiantarMartoba,C1)} = \sqrt{\frac{(1900 - 1996)^2 + (2047 - 2050)^2}{+(2500 - 2500)^2}} = 96,0469$$

$$C_{(SiantarBarat,C1)} = \sqrt{\frac{(1805 - 1996)^2 + (1306 - 2050)^2}{+(1385 - 2500)^2}} = 1353,97$$

$$C_{(SiantarMarihah,C1)} = \sqrt{\frac{(1025 - 1996)^2 + (2050 - 2050)^2}{+(2113 - 2500)^2}} = 1045,28$$

$$C_{(SiantarUtara,C1)} = \sqrt{\frac{(544 - 1996)^2 + (449 - 2050)^2}{+(301 - 2500)^2}} = 3083,36$$

$$C_{(SiantarSelatan,C1)} = \sqrt{\frac{(702 - 1996)^2 + (706 - 2050)^2}{+(760 - 2500)^2}} = 2551,15$$

Cluster Rendah:

$$C_{(SiantarMartoba,C2)} = \sqrt{\frac{(1900 - 87)^2 + (2047 - 144)^2}{+(2500 - 191)^2}} = 3498,55$$

$$C_{(SiantarBarat,C2)} = \sqrt{\frac{(1805 - 87)^2 + (1306 - 144)^2}{+(1385 - 191)^2}} = 2393,2$$

$$C_{(SiantarMarihah,C2)} = \sqrt{\frac{(1025 - 87)^2 + (2050 - 144)^2}{+(2113 - 191)^2}} = 2864,75$$

$$C_{(SiantarUtara,C2)} = \sqrt{\frac{(544 - 87)^2 + (449 - 144)^2}{+(301 - 191)^2}} = 560,334$$

$$C_{(SiantarSelatan,C2)} = \sqrt{\frac{(702 - 87)^2 + (706 - 144)^2}{+(760 - 191)^2}} = 1008,88$$

Begitu seterusnya hingga kota/kab/kec ke 27 baik cluster tinggi maupun rendah. Berikut adalah hasil dari perhitungan jarak cluster:

Tabel 3. Hasil Iterasi 1

No	Kota/Kab/Kec	Tinggi	Rendah	Jarak	C1	C2
1	Siantar Martoba	96,0469	3498,55	96,0469	1	
2	Siantar Barat	1353,97	2393,2	1353,97	1	
3	Siantar Marihat	1045,28	2864,75	1045,28	1	
4	Siantar Utara	3083,36	560,334	560,334		1
5	Siantar Selatan	2551,15	1008,88	1008,88		1

No	Kota/Kab/Kec	Tinggi	Rendah	Jarak	C1	C2
6	Siantar Marimbun	1999,38	1564,41	1564,41		1
7	Siantar Sitalasari	2227,63	1326,95	1326,95		1
8	Raya	2593,99	961,019	961,019		1
9	Bandar	2595,12	968,162	968,162		1
10	Dolok Batunanggar	1727,38	1882,65	1727,38	1	
11	Girsang Sipangan Bolon	3409,85	158,436	158,436		1
12	Pematang Bandar	2891,36	691,475	691,475		1
13	Tanah Jawa	876,986	2855,12	876,986	1	
14	Tapian Dolok	1268,52	2292,18	1268,52	1	
15	Bandar Huluan	2259,63	1315,25	1315,25		1
16	Bandar Masilam	1496,7	2074,26	1496,7	1	
17	Bosar Maligas	3550,86	0	0		1
18	Gunung Malela	2474,6	1098,28	1098,28		1
19	Gunung Maligas	3040,46	540,879	540,879		1
20	Hutabayu Raja	1849,55	1851,36	1849,55	1	
21	Jorlang Hataran	3219,39	440,625	440,625		1
22	Pane	2926,46	676,156	676,156		1
23	Raya Kahean	1633,73	1921,89	1633,73	1	
24	Sidamanik	2267,85	1306,69	1306,69		1
25	Silimakuta	2311,84	1267,46	1267,46		1
26	Ujung Padang	1997,67	1584,48	1584,48		1
27	Serbelawan	598,179	3007,66	598,179	1	

Berdasarkan matrik yang didapatkan pada tabel di atas maka didapatkan pengelompokkan sebagai berikut :

C1 = 1,2,3,10,13,14,16,20,23,27

C2 = 4,5,6,7,8,9,11,12,15,17,18,19,21,22,24,25,26

3. Penentuan pusat cluster baru

Setelah didapatkan hasil dari setiap cluster kemudian pusat cluster baru dihitung berdasarkan data member tiap – tiap cluster yang sudah didapatkan memakai rumus yang sesuai dengan pusat member cluster sebagai berikut :

$$C1(\text{Sep}) = (1900+1805+1025+1357+1996+1424+1330+1445+1114+1857)/10 = 1418,2$$

$$C2(\text{Sep}) = (544+702+1051+800+649+691+221+589+872+87+783+492+374+266+917+898+1119)/17 = 650,294$$

Tabel 4. Pusat Cluster Baru

Cluster	September	Oktober	November
C1	1418,2	1378,5	1590,6
C2	650,294	644,294	675,765

Iterasi selanjutnya dilakukan dengan cara yang sama hingga tidak ada perubahan data dalam suatu cluster.

Iterasi 2:

Cluster Tinggi:

$$C_{(\text{SiantarMartoba}, C1)} = \sqrt{\frac{(1900 - 1418,2)^2 + (2047 - 1378,5)^2}{+(2500 - 1590,6)^2}} = 1227,2$$

$$C_{(\text{SiantarBarat}, C1)} = \sqrt{\frac{(1805 - 1418,2)^2 + (1306 - 1378,5)^2}{+(1385 - 1590,6)^2}} = 444,007$$

$$C_{(\text{SiantarMarihata}, C1)} = \sqrt{\frac{(1025 - 1418,2)^2 + (2050 - 1378,5)^2}{+(2113 - 1590,6)^2}} = 937,241$$

$$C_{(\text{SiantarUtara}, C1)} = \sqrt{\frac{(544 - 1418,2)^2 + (449 - 1378,5)^2}{+(301 - 1590,6)^2}} = 1814,18$$

$$C_{(\text{SiantarSelatan}, C1)} = \sqrt{\frac{(702 - 1418,2)^2 + (706 - 1378,5)^2}{+(760 - 1590,6)^2}} = 1286,5$$

Cluster Rendah:

$$C_{(\text{SiantarMartoba,C2})} = \sqrt{\frac{(1900 - 650,294)^2 + (2047 - 644,29)^2}{+(2500 - 675,76)^2}} = 2618,62$$

$$C_{(\text{SiantarBarat,C2})} = \sqrt{\frac{(1805 - 650,294)^2 + (1306 - 644,29)^2}{+(1385 - 675,76)^2}} = 1508,05$$

$$C_{(\text{SiantarMarihat,C2})} = \sqrt{\frac{(1025 - 650,294)^2 + (2050 - 644,29)^2}{+(2113 - 675,76)^2}} = 2045,01$$

$$C_{(\text{SiantarUtara,C2})} = \sqrt{\frac{(544 - 650,294)^2 + (449 - 644,29)^2}{+(301 - 675,76)^2}} = 435,76$$

$$C_{(\text{SiantarSelatan,C2})} = \sqrt{\frac{(702 - 650,294)^2 + (706 - 644,29)^2}{+(760 - 675,76)^2}} = 116,519$$

Begitu seterusnya hingga kota/kab/kec ke 27 baik cluster tinggi maupun rendah. Berikut adalah hasil dari iterasi 2:

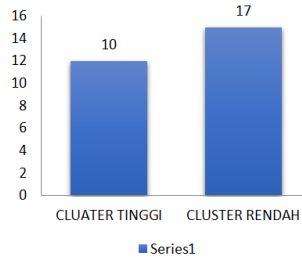
Tabel 5. Hasil Perhitungan Iterasi 2

No	Kota/Kab/Kec	Tinggi	Rendah	Jarak	C1	C2
1	Siantar Martoba	1227,2	2618,62	1227,2	1	
2	Siantar Barat	444,007	1508,05	444,007	1	
3	Siantar Marihat	937,241	2045,01	937,241	1	
4	Siantar Utara	1814,18	435,76	435,76		1
5	Siantar Selatan	1286,5	116,519	116,519		1
6	Siantar Marimbun	734,19	676,278	676,278		1
7	Siantar Sitalasari	981,092	475,92	475,92		1
8	Raya	1336,77	155,006	155,006		1
9	Bandar	1329,55	72,5996	72,5996		1
10	Dolok Batunanggar	480,763	993,848	480,763	1	
11	Girsang Sipangan Bolon	2148,74	751,66	751,66		1
12	Pematang Bandar	1623,74	231,317	231,317		1
13	Tanah Jawa	632,507	1962,91	632,507	1	
14	Tapian Dolok	69,4827	1404,87	69,4827	1	
15	Bandar Huluan	994,669	424,01	424,01		1
16	Bandar Masilam	229,192	1180,72	229,192	1	
17	Bosar Maligas	2292,37	895,875	895,875		1
18	Gunung Malela	1207,12	204,205	204,205		1
19	Gunung Maligas	1773,82	375,126	375,126		1
20	Hutabayu Raja	642,188	974,563	642,188	1	
21	Jorlang Hataran	1954,07	569,799	569,799		1
22	Pane	1685,3	392,631	392,631		1
23	Raya Kahean	418,333	1052,41	418,333	1	
24	Sidamanik	999,099	410,918	410,918		1
25	Silimakuta	1043,2	372,761	372,761		1
26	Ujung Padang	729,067	691,337	691,337		1
27	Serbelawan	739,873	2124,47	739,873	1	

Berdasarkan tabel diatas hasil dari iterasi 2 sama dengan iterasi 1 maka didapatkan pengelompokkan sebagai berikut :

C1 = 1,2,3,10,13,14,16,20,23,27

C2 = 4,5,6,7,8,9,11,12,15,17,18,19,21,22,24,25,26

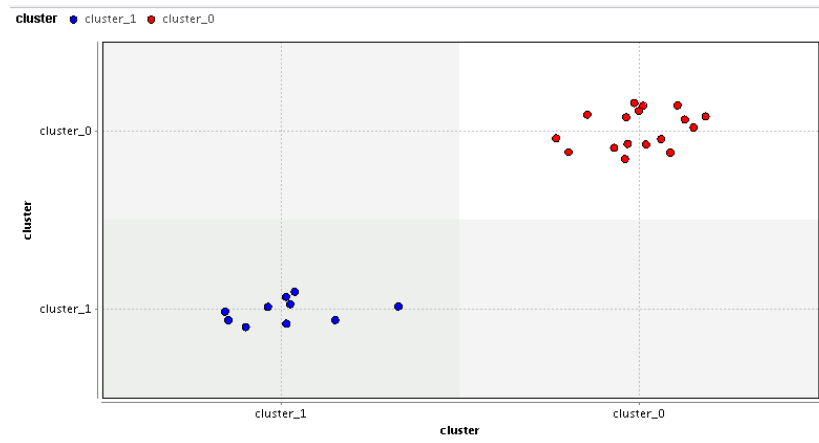


Gambar 1. Grafik Cluster Data

Dari hasil iterasi 2 dapat dilihat bahwa hasil yang diperoleh sama dengan iterasi 2 maka perhitungan dihentikan sampai iterasi 3. Untuk Kota/kab/kec pada cluster C1(tertinggi) yaitu Siantar Martoba, Siantar Barat, Siantar Marihat, Dolok Batunanggar, Tanah Jawa, Tapian Dolok, Bandar Masilam, Hutabayu Raja, Raya Kahean, Serbelawan. Untuk Kota/kab/kec pada cluster C2(rendah) yaitu, Siantar Utara, Siantar Selatan, Siantar Marimbun, Siantar Sitalasari, Raya, Bandar, Girsang Sipangan Bolon, Pematang Bandar, Bandar Hulan, Bosar Maligas, Gunung Malela, Gunung Maligas, Jorlang Hataran, Pane, Sidamanik, Silimakuta, Ujung Padang.

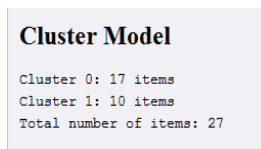
3.2 Pengujian Menggunakan Rapidminer

Setelah melakukan proses sistem dan pengolahan data, maka akan muncul hasil dari *rapidminer* sebagai berikut:



Gambar 2. Grafik Hail Cluster

Berdasarkan gambar diatas dapat dilihat bahwa cluster 0 adalah cluster rendah dan cluster 1 adalah cluster tinggi. Dari gambar diatas juga dapat dilihat bahwa terdapat 10 alternatif yang berada pada cluster tinggi dan 17 alternatif pada cluster rendah seperti yang tertera pada gambar diatas.



Gambar 3. Cluster Model

Berdasarkan dari penjelasan diatas mengenai tahap-tahap penggunaan serta hasil yang telah ditampilkan maka selanjutnya membahas mengenai keterkaitan dari hasil yang didapat antara perhitungan manual algoritma dengan hasil yang ditampilkan oleh *tools rapidminer*. Data di atas akan diuji dengan sampel sebanyak 27 data dan diselesaikan dengan menggunakan *tools rapidminer* sehingga diperoleh hasil cluster seperti pada gambar 4 berikut :

Row No.	Hub	id	cluster	Sep	Okt	Nov
1	Siantar Martoba	1	cluster_1	1900	2047	2500
2	Siantar Barat	2	cluster_1	1805	1306	1385
3	Siantar Marihat	3	cluster_1	1025	2050	2113
4	Siantar Utara	4	cluster_0	544	449	301
5	Siantar Selatan	5	cluster_0	702	706	760
6	Siantar Marimbun	6	cluster_0	1051	931	1139
7	Siantar Sitalasari	7	cluster_0	800	794	1102
8	Raya	8	cluster_0	649	599	824
9	Bandar	9	cluster_0	691	685	720
10	Dolok Batunanggar	10	cluster_1	1357	1051	1244
11	Girsang Sipangan Bolon	11	cluster_0	221	183	266
12	Pematang Bandar	12	cluster_0	589	502	504
13	Tanah Jawa	13	cluster_1	1996	1561	1772
14	Tapian Dolok	14	cluster_1	1424	1323	1632
15	Bandar Hulan	15	cluster_0	872	939	885
16	Bandar Masilam	16	cluster_1	1330	1277	1405
17	Bosar Maligas	17	cluster_0	87	144	191
18	Gunung Malela	18	cluster_0	783	774	761

Gambar 4. Hasil Cluster

Pada gambar diatas dapat dilihat bahwa hasil dari perhitungan manual adalah sama dengan hasil perhitungan menggunakan *tools rapidminer*. Dalam melakukan validasi data perhitungan algoritma harus menghasilkan hasil akhir berupa pengelompokkan dengan 2 *cluster*, serta data yang digunakan merupakan data yang *valid* dan sama dengan yang dipakai pada *tools Rapidminer*. Berikut ditampilkan hasil yang didapatkan dari perhitungan algoritma dan pengujian pada *rapidminer*.

Tabel 6. Perbandingan Hasil Manual dan *RapidMiner*

No	Kota/Kab/Kec	Hasil Perhitungan Manual	Hasil <i>RapidMiner</i>
1	Siantar Martoba	Cluster 1	Cluster 1
2	Siantar Barat	Cluster 1	Cluster 1
3	Siantar Marihat	Cluster 1	Cluster 1
4	Siantar Utara	Cluster 0	Cluster 0
5	Siantar Selatan	Cluster 0	Cluster 0
6	Siantar Marimbun	Cluster 0	Cluster 0
7	Siantar Sitalasari	Cluster 0	Cluster 0
8	Raya	Cluster 0	Cluster 0
9	Bandar	Cluster 0	Cluster 0
10	Dolok Batunanggar	Cluster 1	Cluster 1
11	Girsang Sipangan Bolon	Cluster 0	Cluster 0
12	Pematang Bandar	Cluster 0	Cluster 0
13	Tanah Jawa	Cluster 1	Cluster 1
14	Tapian Dolok	Cluster 1	Cluster 1
15	Bandar Hulan	Cluster 0	Cluster 0
16	Bandar Masilam	Cluster 1	Cluster 1
17	Bosar Maligas	Cluster 0	Cluster 0
18	Gunung Malela	Cluster 0	Cluster 0
19	Gunung Maligas	Cluster 0	Cluster 0
20	Hutabayu Raja	Cluster 1	Cluster 1
21	Jorlang Hataran	Cluster 0	Cluster 0
22	Pane	Cluster 0	Cluster 0
23	Raya Kahean	Cluster 1	Cluster 1
24	Sidamanik	Cluster 0	Cluster 0
25	Silimakuta	Cluster 0	Cluster 0
26	Ujung Padang	Cluster 0	Cluster 0
27	Serbelawan	Cluster 1	Cluster 1

Berdasarkan tabel perbandingan diatas dapat dilihat bahwa hasil perhitungan manual sama dengan hasil dari *rapidminer*. Dimana cluster tinggi sebanyak 10 alternatif yaitu: Siantar Martoba, Siantar Barat, Siantar Marihat, Dolok Batunanggar, Tanah Jawa, Tapian Dolok, Bandar Masilam, Hutabayu Raja, Raya Kahean, Serbelawan. Untuk cluster rendah sebanyak 17 Alternatif yaitu: Siantar Utara, Siantar Selatan, Siantar Marimbun, Siantar Sitalasari, Raya, Bandar,

Girsang Sipangan Bolon, Pematang Bandar, Bandar Hulan, Bosar Maligas, Gunung Malela, Gunung Maligas, Jorlang Hataran, Pane, Sidamanik, Silimakuta, Ujung Padang.

4. KESIMPULAN

- a) Penerapan Data Mining Untuk Pengelompokan Minat Konsumen Terhadap Pengguna Jasa Pengiriman Pada PT. Jalur Nugraha Ekakurir (JNE) Pematangsiantar menjadi pertimbangan pengambil keputusan (decision maker) dalam menentukan minat konsumen pada PT. JNE.
- b) Penelitian ini menggunakan 2 Cluster lokasi dan 2 Iterasi.
- c) Setelah didapatkan hasil dari setiap cluster kemudian pusat cluster baru dihitung berdasarkan data member tiap – tiap cluster yang sudah didapatkan
- d) Iterasi selanjutnya dilakukan dengan cara yang sama hingga tidak ada perubahan data dalam suatu cluster.
- e) Dari hasil setiap iterasi dapat dilihat bahwa hasil yang diperoleh sama dengan cluster tinggi sebanyak 12 alternatif yaitu: Siantar Barat, Siantar Marihat, Siantar Selatan, Siantar Marimbun, Raya, Girsang Sipangan Bolon, Pematang Bandar, Tanah Jawa, Jorlang Hataran, Sidamanik, Silimakuta dan Ujung Padang. Untuk cluster rendah sebanyak 15 Alternatif yaitu: Siantar Martoba, Siantar Utara, Siantar Sitalasari, Bandar, Dolok Batunanggar, Tapian Dolok, Bandar Hulan, Bandar Masilam, Bosar Maligas, Gunung Malela, Gunung Maligas, Hutabayu Raja, Pane, Raya Kahean, dan Siantar.

REFERENCES

- [1] S. F. Rodiyansyah, “Naïve Bayes Classification Untuk Penentuan Kelayakan Donor Darah,” Pp. 156–159, 2014.
- [2] M. G. Sadewo, A. P. Windarto, And D. Hartama, “Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering,” *Infotekjar (Jurnal Nas. Inform. Dan Teknol. Jaringan)*, Vol. 2, No. 1, Pp. 60–67, 2017.
- [3] I. Parlina, A. P. Windarto, A. Wanto, And M. R. Lubis, “Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Assessment Center Untuk Clustering Program Sdp,” *Cess (Journal Comput. Eng. Syst. Sci., Vol. 3, No. 1, Pp. 87–93, 2018.*
- [4] F. S. Jumeilah And D. Pratama, “Klasterisasi Penduduk Lanjut Usia Sumatera Selatan Menggunakan Algoritma K-Modes,” *J. Tam (Technology Accept. Model., Vol. 8, No. 2, Pp. 85–89, 2017.*
- [5] R. Yanto, “Implementasi Data Mining Estimasi Ketersediaan Lahan Pembuangan Sampah Menggunakan Algoritma Regresi Linear,” *J. Resti (Rekayasa Sist. Dan Teknol. Informasi)*, Vol. 2, No. 1, Pp. 361–366, 2018.
- [6] M. Nasution, “Implementasi Data Mining Rough Set Dalam Menentukan Tingkat Kerusakan Alat Dan Bahan Kimia (Studi Kasus Di Laboratorium Resep Smk S-16 Farmasi Bengkulu),” *J. Ilm. Amik Labuhan Batu*, Vol. 2, No. 1, 2014.
- [7] N. Butarbutar, A. P. Windarto, D. Hartama, And Solikhun, “Komparasi Kinerja Algoritma Fuzzy C-Means Dan K-Means Dalam Pengelompokan Data Siswa Berdasarkan Prestasi Nilai akademik Siswa,” *Jurasik (Jurnal Ris. Sist. Inf. Tek. Inform., Vol. 1, No. 2012, 2016.*
- [8] D. S. Kusumo, M. A. Bijaksana, And D. Darmantoro, “Data Mining Dengan Algoritma Apriori Pada Rdbms Oracle,” *J. Penelit. Dan Pengemb. Telekomun., Vol. 8, Pp. 1–5, 2014.*
- [9] S. R. S. Iskandar, Pilipus Tarigan, “Penerapan Data Mining Dalam Surat Jalan Transportasi Mobil Angkutan Menggunakan Metode Clustering,” *J. Pelita Inform., Vol. 17, No. 4, Pp. 460–464, 2018.*
- [10] S. Haryati, A. Sudarsono, And E. Suryana, “Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu),” *J. Media Infotama*, Vol. 11, No. 2, Pp. 130–138, 2015.