

CLUSTERING KESETIAAN PELANGGAN E-RITEL DENGAN MODEL RFM (RECENCY, FREQUENCY, MONETARY) DAN K-MEANS

Tommi Alfian Armawan Sandi¹; Mugi Raharjo²; Jordy Lasmana Putra³; Ridwan⁴

¹Program Studi Ilmu Komputer
STMIK Nusa Mandiri Jakarta
www.nusamandiri.ac.id
alfian.armawan@gmail.com

²Program Studi Ilmu Komputer
STMIK Nusa Mandiri Jakarta
www.nusamandiri.ac.id
mugimou@gmail.com

³Program Studi Ilmu Komputer
STMIK Nusa Mandiri Jakarta
www.nusamandiri.ac.id
balakusalo@gmail.com

⁴Program Studi Ilmu Komputer
STMIK Nusa Mandiri Jakarta
www.nusamandiri.ac.id
ridwans70@gmail.com



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-NonKomersial 4.0 Internasional.

Abstract – *Business is an activity that never stops, everything that becomes a business opportunity can be a promising business for business people, the more advanced and developing business world, making some of the businessmen out of business, many factors that make them troubled in maintaining their business, including customer. This study aims to determine potential and loyal customers to business actors, potential customers determined by customer segmentation. The RFM (Recency, Frequency, Monetary) model is used to find attributes suitable for customer segmentation and pass clusters using the K-Means algorithm, the model issued by potential K-Means customers has a large frequency value. Use the Davies bouldin index to help with the level of accuracy in the cluster data.*

Inti sari – Bisnis merupakan kegiatan yang tidak pernah berhenti, segala sesuatu yang menjadi peluang usaha pun dapat dijadikan bisnis yang menjanjikan kepada pelaku bisnis, semakin maju dan berkembangnya dunia usaha, membuat sebagian dari pebisnis gulung tikar, banyak faktor yang membuat mereka kesusahan dalam mempertahankan bisnisnya, diantaranya adalah pelanggan. Penelitian ini bertujuan untuk menentukan pelanggan potensial dan loyal

kepada pelaku usaha, pelanggan yang potensial ditentukan dengan segmentasi pelanggan. Model RFM (Recency, Frequency, Monetary) digunakan untuk mencari atribut yang cocok untuk segmentasi pelanggan dan melakukan klustering menggunakan algoritma K-Means, model yang di keluarkan oleh K-Means pelanggan yang potensial memiliki nilai frekuensi yang besar. Menggunakan Davies bouldin index untuk membantu tingkat akurasi pada data kluster.

Kata Kunci : Model RFM, K-Means, Davies bouldin index

PENDAHULUAN

Bisnis merupakan kegiatan yang tidak pernah berhenti dalam kehidupan, segala sesuatu yang memiliki peluang usahapun dapat dijadikan bisnis yang bisa memberikan keuntungan kepada pelaku bisnis. Semakin banyaknya pelaku bisnis saat ini berdampak pada persaingan yang cukup ketat. (Reardon, Timmer, & Minten, 2012)

Dengan kemajuan teknologi internet saat ini memunculkan bisnis-bisnis yang baru dalam hal transaksi maupun prinsip kerjanya. Salah satu jenis implementasi online dalam

bisnis adalah *e-commerce* (Iramawati, 2011) Menurut survey yang dikeluarkan oleh statista.com tentang transaksi belanja online pada tahun 2017, diperkirakan 1,66 miliar orang di seluruh dunia membeli barang secara online. Selama tahun yang sama, penjualan e-ritel global berjumlah 2,3 triliun dolar AS dan proyeksi menunjukkan pertumbuhan hingga 4,48 triliun dolar AS pada tahun 2021. Di Asia Pasifik, penjualan e-ritel menyumbang 12,1 persen dari penjualan ritel pada tahun 2016 tetapi hanya untuk 1,8 persen dari penjualan ritel di Timur Tengah dan Afrika (Duncan, 2018), dari survey tersebut dapat disimpulkan bahwa saat ini transaksi online mengalami peningkatan sejalan dengan modernisasi pada ritel. Kunci utama supaya ritel dapat berjalan adalah kesetiaan pelanggan pada produk-produk ritel tersebut.

Pada penelitian sebelumnya komparasi metode *clustering k-means* dan *k-medoids* dengan model *fuzzy RFM* untuk pengelompokan pelanggan. Permasalahan yang terjadi pada pelaku usaha disebabkan karena mereka mengalami kesulitan dalam menentukan pelanggan yang potensial. (Muningsih, 2018), membandingkan dan mengkomparasi dua metode *clustering* dengan cara mengelompokkan pelanggan menjadi 3 cluster sesuai karakteristiknya, metode K-means lebih baik dari metode K-medoids dengan akurasi 90,47% (Muningsih, 2018). Dapat disimpulkan terdapat pola dari mutu kluster dengan pemilihan atribut menggunakan metode RFM untuk mengkluster pelanggan yang lebih efisien dan akurat dengan algoritma *Fuzzy C-means* untuk memilih pelanggan yang potensial dan loyal (Sudriyanto, 2017).

Dengan membandingkan penelitian yang sebelumnya penulis akan membahas tentang *clustering kesetiaan pelanggan* dengan model *RFM (Recency, Frequency, Monetary)* dan *K-Means*. Tujuan dari penelitian ini adalah mengelompokkan pelanggan menurut peluang yang akan didapat dengan memerhatikan tiga faktor pelanggan yaitu, pelanggan sangat potensial, pelanggan potensial dan pelanggan yang kurang potensial.

BAHAN DAN METODE

Dalam penelitian ini akan dilakukan beberapa tahapan eksperimen menggunakan tahapan proses penelitian berikut :

1. Pengumpulan Data

Data yang dilakukan pada penelitian ini bersumber pada *database UCI Dataset Ritel Online*. Data yang digunakan adalah sebagai berikut :

- 1) Data pelanggan yang melakukan transaksi.

- 2) Data riwayat transaksi dengan atribut yang dibutuhkan oleh RFM, Terdiri dari:
 - a) Waktu transaksi terakhir
 - b) jumlah nominal transaksi
 - c) jumlah frekuensi transaksi

2. Praposes Data

Data yang sudah diperoleh dari data transaksi pelanggan akan melalui praproses data yang meliputi (Sudriyanto, 2017) :

- 1) *Data cleaning and integration*, yaitu proses pembersihan data untuk menghilangkan baris data dengan nilai kosong atau data yang tidak valid.
- 2) *Data selection*, yaitu pemilihan data berdasarkan atribut yang disesuaikan dengan model RFM, meliputi : atribut waktu transaksi, jumlah frekuensi transaksi dan jumlah nominal transaksi untuk setiap pelanggan.
- 3) *Data preprocessing*, yaitu persiapan data dengan cara mereduksi kolom data yang tidak sesuai dengan kebutuhan dalam melakukan proses klusterisasi.
- 4) *Transformation*, yaitu transformasi data ke dalam bentuk terukur sehingga dapat digunakan sebagai atribut untuk proses klusterisasi.

3. Perhitungan Nilai RFM

Model RFM (*Recency, Frequency dan Monetary*) digunakan untuk menentukan variabel mengukur pembelian produk oleh pelanggan. Variabel dapat menentukan sebagai kebaruan, frekuensi dan moneter (Putu et al., 2015). Model RFM ini merupakan metode yang sudah lama dan populer untuk mengukur hubungan dengan pelanggan (Hardiani & Hartanto, 2017). Terdapat tiga kriteria dalam model RFM yaitu *recency, frequency, dan monetary*.

- a. *Recency* adalah rentang waktu (hari, bulan, tahun) dari transaksi akhir sampai saat ini oleh pelanggan membeli.
- b. *Frequency* atau frekuensi adalah total transaksi atau transaksi rata-rata dalam satu kali periode.
- c. *Monetary* atau moneter adalah biaya rata-rata total pelanggan di sekali waktu.

Data yang sudah dilakukan pengecekan atribut-atribut praproses maka selanjutnya dilakukan proses klusterisasi.

4. Metode yang diusulkan

Algoritma K-Means

K-Means Cluster Analysis merupakan salah satu metode cluster analysis non hirarki yang

berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain (Mara & Satyahadewi, 2013) .

Data clustering menggunakan metode K-Means ini secara umum dilakukan dengan algoritma dasar sebagai berikut (Agusta, 2007):

1. Tentukan jumlah cluster
2. Alokasikan data ke dalam cluster secara random
3. Hitung *centroid*/rata-rata dari data yang ada di masing-masing cluster
4. Alokasikan masing-masing data ke *centroid*/rata-rata terdekat
5. Kembali ke Step 3, apabila masih ada data yang berpindah cluster atau apabila perubahan nilai *centroid*, ada yang di atas nilai *threshold* yang ditentukan atau apabila perubahan nilai pada objective function yang digunakan di atas nilai *threshold* yang ditentukan.

Perhitungan jarak antar titik dengan menggunakan *Euclidean distance*. Yaitu :

$$d(x^j, c^j) = \sqrt{\sum_{j=1}^n (x_i - c_i)^2} \dots\dots\dots (1)$$

Dimana :
 d = Jarak
 j = banyaknya data
 c = centroid
 x = data

5. Evaluasi dan Validasi Hasil

Evaluasi dilakukan menggunakan Rapidminer 9.0 dengan melihat hasil akhir yang dihitung secara manual dan otomatis. Dengan validasi menggunakan model davies bouldin index (Dbi). Menurut (S, Bernad Jumadi Dehotman, 2018) Davies-Bouldin Index merupakan salah satu metode yang digunakan untuk mengukur validitas cluster pada suatu metode pengelompokan, kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap titik pusat cluster dari cluster yang diikuti. Sedangkan separasi didasarkan pada jarak antar titik pusat cluster terhadap clusternya.

Tahapan dari perhitungan DaviesBouldin Index adalah sebagai berikut:

1. Sum of Square Within-cluster (SSW)

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \dots\dots\dots (2)$$

2. Sum of Square Between-cluster (SSB)

$$SSB_{i,j} = d(c_i, c_j) \dots\dots\dots (3)$$

3. Ratio

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \dots\dots\dots (4)$$

4. Davies Bouldin Index

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \dots\dots\dots (5)$$

HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Data yang di uji bersumber dari UCI Reporsitory Retail Online, dalam data yang diambil terdapat atribut-atribut seperti invoice number, stock code, description, Quantity, invoice date, unit price, total, customerID, country. Periode yang digunakan dari data tersebut 1 desember 2010 sampai 18 januari 2011.

2. Praproses data

Pada tahapan ini data yang telah dikumpulkan akan masuk pada tahap pemilihan atribut, atribut yang akan dipilih merupakan atribut yang digunakan untuk RFM.

Tabel 1. Tabel Pelanggan

No	Nama Atribut	Keterangan
1	Invoice No	Nomor faktur transaksi pelanggan
2	Stock code	Kode persediaan barang
3	Description	Nama barang yang dibeli pelanggan
4	Quantity	Jumlah satuan barang yang dibeli pelanggan
5	Invoice date	Tanggal pada saat pelanggan melakukan transaksi
6	Unit price	Harga satuan barang
7	Total	Jumlah harga satuan dikali quantity
8	Customer ID	Kode pelanggan yang melakukan transaksi
9	Country	Negara terjadinya transaksi

Sumber: (Sandi, Raharjo, Putra, & Ridwan, 2019)

Setelah melakukan seleksi atribut maka atribut yang tersisa hanya sebagai berikut :

Tabel 2. Pemilihan Atribut RFM

No	Nama Atribut	Keterangan
1	Invoice date	Tanggal pada saat pelanggan melakukan

		transaksi
2	Costomer ID	Kode pelanggan yang melakukan transaksi
3	Total	Jumlah harga yang dibeli pelanggan

Sumber : (Sandi et al., 2019)

3. Penilaian pelanggan berbasis model RFM

Pada tahap ini, atribut yang sudah di praproses disesuaikan dengan kriteria pada model RFM , kriteria R (recency) Membutuhkan atribut yang menunjukkan adanya rentang waktu transaksi terakhir pelanggan dengan periode analisis, sehingga atribut yang dibutuhkan adalah atribut tanggal transaksi. kriteria F (Frekuensi) Membutuhkan atribut yang merepresentasikan berapa kalipelanggan melakukan transaksi. Kriteria ini ini dapat dilihat dari berapa banyak pelanggan dengan nama yang sama muncul dalam data transaksi. Atribut yang dibutuhkan adalah atribut tanggal transaksi yang dihitung jumlahnya, kriteria M (monetary) Membutuhkan atribut yang berhubungan dengan harga yang telah dihabiskan pelanggan selama melakukan transaksi, sehingga atribut yang dibutuhkan adalah atribut harga total yang harus dibayar pelanggan.

Berikut adalah data yang sudah melalui tahapan praproses dan selanjutnya dilakukan proses klasterisasi.

Tabel 3 . Hasil dari transformasi data ke model RFM (Recency, Frequency, Monetary)

No	CS ID	R	F	M
1	12346	2	0	0
2	12347	29	315	475.39
3	12348	6	601	227.44
4	12350	17	197	334.4
5	12352	15	98	296.5
6	12356	36	1216	2271.62
7	12359	96	609	2386.41
8	12362	27	229	479.1
9	12365	23	173	320.69
10	12372	20	224	399.22
11	12373	14	197	364.6
12	12377	34	340	626.6
13	12383	32	587	639.91
14	12386	2	140	143
.....
449	17841	67	196	533.58
450	17850	84	474	1499.34
451	17873	9	129	215.15
452	17897	36	66	140.39

453	17905	23	72	201.75
454	17908	58	173	243.28
455	17920	81	220	514.41
456	17924	2	60	279
457	17951	5	118	295.5
458	17968	85	160	277.35
459	18074	13	190	489.6
460	18085	9	78	303.9
461	18144	3	95	165.05
462	18229	7	44	344.2

Sumber : (Sandi et al., 2019)

4. Proses Klustering menggunakan algoritma K-Means

Tahapan dalam pemakaian algoritma K-means adalah menentukan banyaknya cluster, cluster yang akan dibuat sebanyak 3 cluster. Banyaknya cluster harus lebih kecil dari pada banyaknya data. Lalu, menentukan secara acak nilai centroid awal dari data yang sudah dimiliki.

Tabel 4. Cluster Awal

	Recency	Frequency	Monetary
C1	20	224	399.22
C2	7	73	120.95
C3	3	116	357

Sumber : (Sandi et al., 2019)

Langkah selanjutnya adalah memasukan kedalam rumus Euclidean distance space, untuk menghitung jarak data ke pusat centroid.

$$d(x^j, c^j) = \sqrt{\sum_{i=1}^n (x_i - c_i)^2} \dots\dots\dots (6)$$

- Dari data pertama ke pusat cluster 1

$$d(x^1, c^1) = \sqrt{(2 - 20)^2 + (0 - 224)^2 + (0 - 399.22)^2} = 458.1229$$

Dari hasil perhitungan diatas didapat hasil bahwa jarak data pertama dengan cluster pertama adalah 458.1229

- Dari data pertama ke pusat cluster 2

$$d(x^1, c^2) = \sqrt{(2 - 7)^2 + (0 - 73)^2 + (0 - 120.95)^2} = 141.3609$$

Dari hasil perhitungan diatas didapat hasil bahwa jarak data pertama dengan cluster kedua adalah 141.3609

Dari data pertama ke pusat cluster 3

$$d(x^1, c^2) = \sqrt{(2 - 3)^2 + (0 - 116)^2 + (0 - 357)^2} = 375.3745$$

Dari hasil perhitungan diatas didapat hasil bahwa jarak data pertama dengan cluster ketiga adalah 375.3745

Selanjutnya kelompokan data yang memiliki jarak terpendek pada tiap clusternya. Berdasarkan hasil perhitungan ketiga cluster tersebut dapat disimpulkan cluster 2 adalah jarak yang paling terdekat dengan centroid, maka beri tanda pada C2. Untuk hasil perhitungan lengkapnya dapat dilihat pada Tabel 5.

Tabel 5. Jarak data perama dengan pusat cluster pada iterasi 1

No	Customer ID	Recency	Frequency	Monetary	C1	C2	C3	c1	c2	c3
1	12346	2	0	0	458.1229	141.3609	375.3745	X		
2	12347	29	315	475.39	119.0121	429.7391	233.009	X		
3	12348	6	601	227.44	414.5279	538.6326	502.0157	X		
4	12350	17	197	334.4	70.28252	247.0565	85.25116	X		
5	12352	15	98	296.5	162.6419	177.5016	64.25146			X
6	12356	36	1216	2271.62	2119.01	2435.708	2208.361	X		
7	12359	96	609	2386.41	2025.568	2329.705	2090.503	X		
8	12362	27	229	479.1	80.34186	391.1616	168.0875	X		
9	12365	23	173	320.69	93.68544	223.9466	70.4799			X
10	12372	20	224	399.22	0	316.8662	117.1987	X		
...
...
455	17920	81	220	514.41	130.406	426.4924	204.1517	X		
456	17924	2	60	279	204.1393	158.6625	96.02604			X
457	17951	5	118	295.5	149.0598	180.2684	61.56501			X
458	17968	85	160	277.35	152.2278	195.228	122.4913			X
459	18074	13	190	489.6	96.81707	386.8176	152.18	X		
460	18085	9	78	303.9	174.708	183.0292	65.57141			X
461	18144	3	95	165.05	267.891	49.44502	193.0953		X	
462	18229	7	44	344.2	188.6696	225.1257	73.23824			X

Sumber : (Sandi et al., 2019)

Setelah semua data ditempatkan ke dalam cluster yang terdekat, kemudian hitung kembali pusat cluster yang baru berdasarkan rata-rata anggota yang ada pada cluster tersebut.

$$\text{Rata-rata Recency pada cluster 1} = \frac{29 + 6 + 17 + 36 + 96 + 27 + \dots + 81 + 13}{241}$$

$$= \frac{10746}{241} = 44.58921$$

$$\text{Rata-rata Recency pada cluster 2} = \frac{2 + 2 + 5 + 11 + 7 + \dots + 36 + 23 + 3}{150}$$

$$= \frac{1677}{150} = 11.18$$

$$\text{Rata-rata Recency pada cluster 3} = \frac{15 + 23 + 19 + 13 + 15 + 16 + \dots + 58 + 85}{71}$$

$$= \frac{1251}{71} = 17.61972$$

$$\text{Rata-rata Frequency pada cluster 1} = \frac{315 + 601 + 197 + 1216 + 609 + \dots + 220}{241}$$

$$= \frac{168791}{241} = 700.3776$$

$$\text{Rata-rata Frequency pada cluster 2} = \frac{140 + 11 + 94 + 145 + 104 + 194 + \dots + 95}{150}$$

$$= \frac{9977}{150} = 66.51333$$

$$\text{Rata-rata Frequency pada cluster 3} = \frac{98 + 173 + 103 + 113 + 105 + \dots + 78 + 44}{71}$$

$$= \frac{9072}{71} = 127.7746$$

$$\begin{aligned} &\text{Rata-rata Monetary pada cluster 1} \\ &= \frac{475.39 + 227.44 + 334.4 + \dots + 489.6}{241} \\ &= \frac{300922.4}{241} = 1248.641 \end{aligned}$$

$$\begin{aligned} &\text{Rata-rata Monetary pada cluster 2} \\ &= \frac{143 + 84.3 + 225.6 + \dots + 165.05}{150} \\ &= \frac{15987.39}{150} = 105.9826 \end{aligned}$$

$$\begin{aligned} &\text{Rata-rata Monetary pada cluster 3} \\ &= \frac{296.5 + 320.69 + 291.49 + \dots + 344.2}{71} \\ &= \frac{21742.69}{71} = 306.2351 \end{aligned}$$

Kemudian, hasil dari perhitungan rata-rata cluster 1,2,3 menjadikan suatu centroid baru Langkah berikutnya adalah mengulang perhitungan pada tiap cluster sampai didapatkan hasil akhir yang tidak berubah/berpindah. Pada penelitian ini hasil akhir berada pada iterasi 7. Hasil dan pola yang didapat pada tabel 6 dan tabel 7.

Tabel 6. Centroid baru

	Recency	Frequency	Monetary
C1	23.5625	1977	16.54351
C2	25.375	11475.5	101.4994
C3	805.6875	80467.5	644.4851

Sumber : (Sandi et al., 2019)

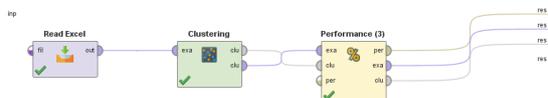
Tabel 7. Hasil dan pola terakhir jarak antara centroid dan pusat cluster

No	Customer ID	Recency	Frequency	Monetary	C1	C2	C3	c1	c2	c3
1	12346	2	0	0	82444.5	11475.5	2			X
2	12347	29	315	475.39	82130.88	11170.66	571.0181			X
3	12348	6	601	227.44	81843.82	10876.88	642.6243			X
4	12350	17	197	334.4	82248.18	11283.47	388.486			X
5	12352	15	98	296.5	82347.03	11381.37	312.636			X
6	12356	36	1216	2271.62	81260.26	10508.04	2576.86			X
7	12359	96	609	2386.41	81870.32	11125.86	2464.762			X
8	12362	27	229	479.1	82216.9	11256.73	531.7018			X
9	12365	23	173	320.69	82272.13	11307.07	365.1028			X
10	12372	20	224	399.22	82221.47	11258.6	458.2059			X
...
...
455	17920	81	220	514.41	82226.13	11267.53	565.3129			X
456	17924	2	60	279	82384.98	11418.91	285.3857			X
457	17951	5	118	295.5	82327.03	11361.34	318.2283			X
458	17968	85	160	277.35	82284.99	11319.21	331.2824			X
459	18074	13	190	489.6	82255.96	11296.12	525.3353			X
460	18085	9	78	303.9	82367.06	11401.55	313.8793			X
461	18144	3	95	165.05	82349.67	11381.7	190.4613			X
462	18229	7	44	344.2	82401.22	11436.68	347.0715			X

Sumber : (Sandi et al., 2019)

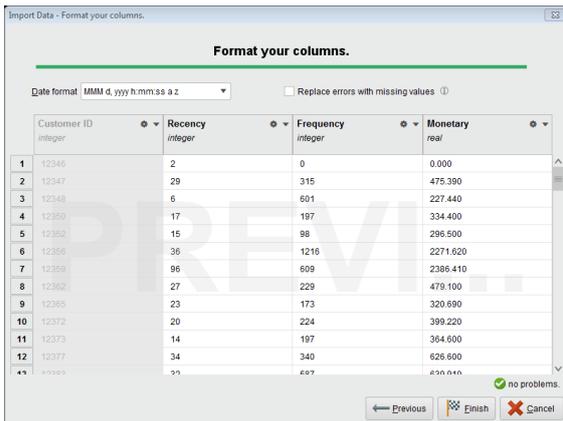
5. Evaluasi dan Validasi Hasil

Pengujian dilakukan dengan jarak kedekatan Euclidean distance menggunakan rapid miner 9.0



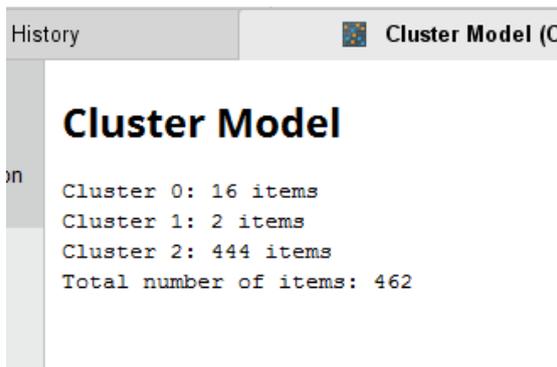
Sumber : (Sandi et al., 2019)

Gambar 1. Tampilan Design Euclidean distance



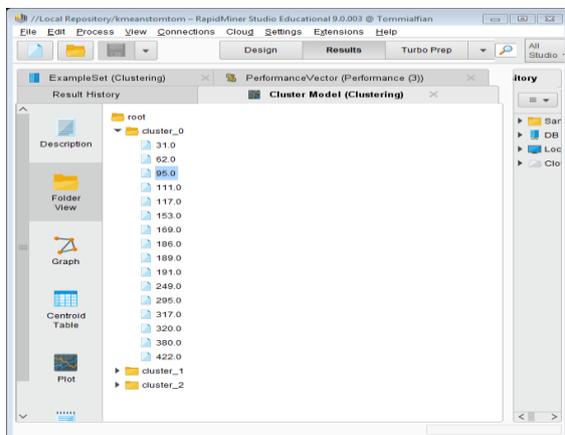
Sumber : (Sandi et al., 2019)
Gambar 2. Tampilan memilih atribut

Pada cluster mode terbentuk beberapa cluster, yaitu cluster 0 dengan jumlah 16 item, cluster 1 2 items dan cluster2 444 item dengan total data 462 item, baik manual maupun aplikasi data yang sajikan sudah sama,



Sumber : (Sandi et al., 2019)
Gambar 3. Cluster model

Cluster model dapat juga dilihat secara mendetail dengan menu folder view, pada folder view terdapat root folder yang berisi cluster-cluster yang berbeda adalah, terdapat melihat anggota pada tiap cluster.



Sumber : (Sandi et al., 2019)
Gambar 4. Tampilan folder view

Data yang dikelompokkan pada cluster 1 berjumlah 16 orang di cluster 2 berjumlah 2 orang dan cluster 3 berjumlah 444 orang. Hasil Analisa pelanggan yang potensial

Davies bouldin index

Jika nilai dbi terkecil maka itulah jumlah cluster yang baik, semakin kecil nilainya maka semakin optimal hasil clusternya, proses clustering dan pengujian davies bouldin dapat dilihat pada tabel 7, seperti yang ditunjukkan oleh tabel didapatkan jumlah cluster 3 yang paling baik. Nilai negatif yang tinggi menunjukkan kinerja yang baik dari indeks. (Hardiani & Hartanto, 2017)

Tabel 7. Perbandingan Hasil Pengujian Davies Bouldin Index

Jumlah cluster	Davies Bouldin Index
2	-0.148
3	-0.728
4	-0.460
5	-0.611
6	-0.644
7	-0.502
8	-0.644
9	-0.562
10	-0.542

Sumber : (Sandi et al., 2019)

KESIMPULAN

Cluster yang terbentuk berdasarkan RFM, semakin besar Nilai R menunjukkan bahwa pelanggan sering melakukan transaksi, semakin besar nilai F menunjukkan pelanggan setia terhadap pelaku usaha dan semakin besar nilai M, menunjukkan bahwa nominal transaksi yang dilakukan semakin besar. Dengan menggunakan DBI dapat memastikan cluster yang digunakan adalah tepat dan benar supaya hasil yang didapat akurat.

REFERENSI

Agusta, Y. P. (2007). K-Means – Penerapan, Permasalahan dan Metode Terkait, 3(Pebruari), 47-60.

Duncan, E. (2018). Online-Shopping and E-Commerce worldwide: Statistics & Facts. Retrieved from <https://www.statista.com/topics/871/online-shopping/>

Hardiani, T., & Hartanto, R. (2017). Segmentasi Nasabah Tabungan Menggunakan Model RFM (Recency , Frequency , Monetary) dan

K-Means Pada Lembaga Keuangan Mikro, (May).

Iramawati, D. (2011). PEMANFAATAN E-COMMERCE DALAM DUNIA BISNIS. *Jurnal Ilmiah Orasi Bisnis, VI*, 95–112.

Mara, M. N., & Satyahadewi, N. (2013). PENGKLASIFIKASIAN KARAKTERISTIK DENGAN METODE K-MEANS CLUSTER ANALYSIS, *02(2)*, 133–136.

Muningsih, E. (2018). KOMPARASI METODE CLUSTERING K-MEANS DAN K-MEDOIDS DENGAN MODEL FUZZY RFM UNTUK PENGELOMPOKAN PELANGGAN, *6(2)*.

Putu, N. I., Yuliari, P., Gede, I. K., Putra, D., Kadek, N. I., & Rusjayanti, D. W. I. (2015). CUSTOMER SEGMENTATION THROUGH FUZZY C-MEANS AND FUZZY RFM METHOD, *78(3)*, 380–385.

Reardon, T., Timmer, C. P., & Minten, B. (2012). Supermarket revolution in Asia and emerging development strategies to include small farmers, *109(31)*.

<https://doi.org/10.1073/pnas.1003160108>

Sandi, T. A. A., Raharjo, M., Putra, J. L., & Ridwan. (2019). *Laporan Akhir Mandiri - Clustering Kesetiaan Pelanggan E-Ritel Dengan Model Rfm (Recency, Frequency, Monetary) Dan K-Means*.

Sudriyanto. (2017). CLUSTERING LOYALITAS PELANGGAN DENGAN METODE RFM (RECENCY, FREQUENCY, MONETARY) DAN FUZZY C-MEANS, 815–822. Retrieved from RFM (Recency, Frequency, Monetary), Fuzzy C-Means, clustering, validitas cluster, Partition Coefficient Index (PCI) dan Xie dan Beni (XBI)