

**CLUSTERING LOYALITAS PELANGGAN DENGAN MODEL RFM
(*REFENCY, FREQUENCY, MONETARY*) DAN METODE K-MEANS
BERBASIS MEDIAN**Ahmad Supriadi¹, Moh. Sukron², Mochammad Faid³
^{1,2,3} Universitas Nurul Jadid, Indonesia**Info Artikel**Riwayat Artikel

Diterima: 20-10-2021

Disetujui: 26-11-2021

Kata KunciRFM (*Recency, Frequency, Monetary*);
K-Means Berbasis Median;
Clustering Loyalitas Pelanggan;

e-mail*

cahaya.supriady04@gmail.com

ABSTRAK

Data loyalitas pelanggan dalam dunia jual beli selalu bertambah seiring dengan berjalannya waktu. Perkembangan teknologi yang sangat pesat dalam dunia bisnis perdagangan menjadi arah perubahan yang sangat signifikan dalam proses transaksi jual beli dengan pemanfaatan teknologi proses transaksi penjualan dapat dijadikan sebuah sekumpulan data besar dan melimpah. Hal ini berpeluang untuk dapat dilakukan proses pengolahan data sehingga akan didapat sebuah informasi yang bermanfaat. Informasi tersebut seperti mengenai pengelompokan data pelanggan yang memiliki loyalitas tinggi sehingga pihak unit usaha terkait dapat memberikan reward pada pelanggan yang memiliki tingkat loyalitas tinggi. Untuk mengetahui hal tersebut, dilakukan pengujian model dengan menggunakan metode k-means dengan pemilihan pusat kluster awal berbasis median, hasil eksperimen dengan model tersebut dapat menghasilkan kualitas kluster yang lebih baik karena dengan pemilihan pusat kluster awal berbasis median tidak terjadi kluster kosong. Kluster yang dihasilkan yaitu pada cluster pertama 133 cluster Kedua 30 dan cluster ketiga 208 sedangkan cluster keempat 129 dengan nilai DBI 0,694. Sedangkan metode k-means dengan pemilihan pusat kluster standart didalam eksperimen ini menggunakan tool rapidminer didapat cluster pertama 310 cluster Kedua 19 dan cluster ketiga 37 sedangkan cluster keempat 134 dengan nilai DBI 0,807.

1. PENDAHULUAN

loyalitas pelanggan adalah "*Loyalty is defined as non random purchase expressed over time by some decision making unit*"[1]. Makna dari definisi diatas adalah parameter titik acuan loyalitas terletak pada sebuah bentuk pola perilaku dari setiap unit untuk mengambil keputusan dalam melakukan transaksi pembelian dengan cara terus menerus pada barang maupun jasa unit usaha yang akan dipilih. Terdapat 3 macam kriteria customer loyal, yang pertama menggunakan produk maupun jasa perusahaan tanpa membandingkannya dengan penawaran perusahaan lain, yang kedua merekomendasikan unit usaha, pelayanan dan produk unit usaha melalui mulut ke mulut pada orang lain serta melakukan tindakan pro aktif memberikan saran produk atau jasa karena unit usaha lain.

Perkembangan teknologi yang sangat pesat dalam dunia bisnis perdagangan menjadi arah perubahan yang sangat signifikan dalam proses transaksi jual beli[3]. Dengan pemanfaatan teknologi proses transaksi penjualan dapat dijadikan sebuah sekumpulan data besar dan melimpah. Hal ini berpeluang untuk dapat dilakukan proses pengolahan data sehingga akan didapat sebuah informasi yang bermanfaat. Informasi tersebut seperti mengenai

pengelompokan data pelanggan yang memiliki loyalitas tinggi sehingga pihak unit usaha terkait dapat memberikan reward pada pelanggan yang memiliki tingkat loyalitas tinggi.

Penelitian sebelumnya ialah tentang Segmentasi Nasabah Tabungan Menggunakan Model RFM (Recency, Frequency, Monetary) dan K-Means Pada Lembaga Keuangan Mikro[4]. Penelitian ini fokus pada segmentasi nasabah setelah melakukan segmentasi didapat 3 hasil cluster yang paling tepat dan optimal. pada pengujian cluster dalam penelitian ini menggunakan Davies Bouldin Index (DBI). Cluster pertama terdiri dari 100 nasabah. RFM score antara 112-255 termasuk kelompok occasional customer. Cluster ke dua terdiri dari 69 nasabah, memiliki RFM score antara 441-544 dan termasuk kelompok superstar. Dan Cluster ke tiga terdiri dari 79 nasabah, memiliki RFM score antara 313-435 dan termasuk kelompok typical customer. Salah satu metode pengolahan data untuk mendapatkan informasi atau pengetahuan baru ialah metode Clustering. Klustering adalah proses pengelompokan data menjadi beberapa cluster sehingga data dalam satu cluster memiliki tingkat kemiripan maksimum dan data antar cluster memiliki kemiripan minimum [5]. Klustering mengacu pada pembagian data ke dalam kelompok objek yang sama. Setiap kelompok, terdiri dari elemen-elemen yang mirip satu sama lain dan berbeda dengan objek dalam kelompok lain. Klustering dalam data mining berguna untuk menemukan pola distribusi di dalam sebuah dataset yang berguna untuk proses analisa data, suatu bentuk pemodelan data yang bersumber pada ilmu matematika dan statistik. Dari perspektif pembelajaran mesin, klustering banyak digunakan untuk mengenali pola yang tersembunyi, model klustering merupakan pembelajaran tanpa pengawasan, dan sistem yang dihasilkan merupakan konsep data[6]. Kesamaan objek biasanya diperoleh dari kedekatan nilai atribut yang menggambarkan objek data, sedangkan objek data biasanya direpresentasikan sebagai titik dalam ruang multidimensi. Klustering dapat mengidentifikasi daerah padat, pola distribusi keseluruhan dan keterkaitan antara atribut data. Klustering berguna dalam beberapa analisis eksplorasi pola, pengelompokan, pengambilan keputusan, dan situasi pembelajaran mesin, termasuk data mining, pengambilan dokumen, segmentasi citra, dan klasifikasi pola.

K-means adalah salah satu algoritma klustering yang cukup sederhana dan populer untuk meramalkan dataset kedalam beberapa kluster k. Algoritmanya relatif mudah untuk diimplementasikan serta relatif cepat, mudah disesuaikan serta banyak digunakan [7]. Menyusun data kedalam kluster dengan mengelompokkan data berdasarkan kedekatan jarak antara data dari sekumpulan data yang ada dalam dataset merupakan teknik utama dalam algoritma ini. Langkah awal dalam algoritma K-means yaitu diawali dengan pembentukan partisi kluster di awal, setelah itu mengelompokkan data berdasarkan kedekatan jarak antar data secara iteratif, partisi kluster ini diperbaiki hingga tidak terjadi perubahan yang signifikan pada partisi kluster[8].

Algoritma K-means merupakan algoritma yang paling populer dan paling banyak digunakan untuk proses klustering data[9], [10]. Popularitasnya dapat dikaitkan dengan beberapa alasan. Pertama, secara konseptual sederhana dan mudah diimplementasikan. Hampir setiap software data mining mencakup pelaksanaan itu. Kedua, serbaguna yaitu hampir setiap aspek dari algoritma (inisialisasi, fungsi jarak, kriteria terminasi, dll) dapat dimodifikasi. Hal ini dibuktikan oleh banyak publikasi selama beberapa tahun terakhir yang memperbaiki metode k-means dalam berbagai cara, akan tetapi di sisi lain k-means memiliki beberapa kelemahan yang signifikan yaitu hasil dari klusterisasi sangat bergantung pada pemilihan pusat kluster awal, sehingga ketika pemilihan pusat kluster awal buruk maka dengan mudah terjebak dalam poor local minima[11],[12]. K-means sangat sensitif terhadap penentuan pusat kluster awal, karena akan mengakibatkan kluster akhir tidak optimal dan bahkan memungkinkan kluster kosong jika tidak ada poin yang dialokasikan kedalam kluster selama iterasi berlangsung[13]. Untuk alasan yang sama, K-means sangat sensitif terhadap pemilihan pusat kluster awal. Efek samping dari inisialisasi yang tidak tepat adalah memungkinkan adanya kluster kosong, konvergensi lambat, dan kesempatan yang lebih tinggi untuk terjebak dalam poor local minima[14]. Oleh karena itu, cukup penting bagi K-means

untuk memiliki pusat kluster awal yang baik.

Pada penelitian ini digunakan algoritma RFM dalam penentuan atribut dan algoritma K-Means dengan memanfaatkan perhitungan nilai median dalam penentuan kluster untuk pengelompokan tingkat loyalitas pelanggan sehingga mendapatkan hasil pengelompokan terbaik.

2. RUMUSAN MASALAH

Berdasarkan latar belakang diatas yang telah dipaparkan, maka permasalahan yang ditimbulkan adalah:

Bagaimana proses optimalisasi metode K-Means dalam mengkluster data tingkat loyalitas pelanggan dengan menggunakan perhitungan algoritma median dalam penentuan pusat kluster awal ?

3. TUJUAN PENELITIAN

Berdasarkan rumusan masalah yang telah diuraikan diatas, maka penelitian ini bertujuan untuk :

- a. Melakukan proses klustering data tingkat loyalitas pelanggan menggunakan metode K-Means dengan optimalisasi algoritma median dalam penentuan pusat kluster awal supaya dapat menghasilkan hasil kluster terbaik sehingga dijadikan acuan atau bahan pertimbangan bagi pihak pemilik unit usaha untuk memberikan reward pada pelanggan yang memiliki loyalitas tinggi.
- b. Penggunaan optimalisasi Algoritma Median dalam penentuan pusat kluster awal pada model K-Means supaya dapat menghasilkan hasil klusterisasi yang terbaik.

4. PENELITIAN TERKAIT

Penelitian dilakukan Sudriyanto[3]. Pada penelitian ini menggunakan Fuzzy C-Means dalam menentukan kluster dan RFM untuk proses pemilihan atribut. Adapun hasil dari penelitian ini adalah mengukur nilai validitas menggunakan FCM dengan pangkat dua 0,8156 dan pangkat tiga 0,5860 serta menggunakan Partition Coefficient Index (PCI). Dalam penelitian ini nilai yang semakin mendekati 1 memiliki arti bahwa kualitas clusternya semakin baik. Lalu metode pengukuran dengan Xie Beni Index (XBI) dengan nilai FCM pangkat dua 0,0069 dan pangkat tiga 0,0632. Ini bernilai index yang semakin kecil mempunyai arti hasil pengelompokan semakin baik.

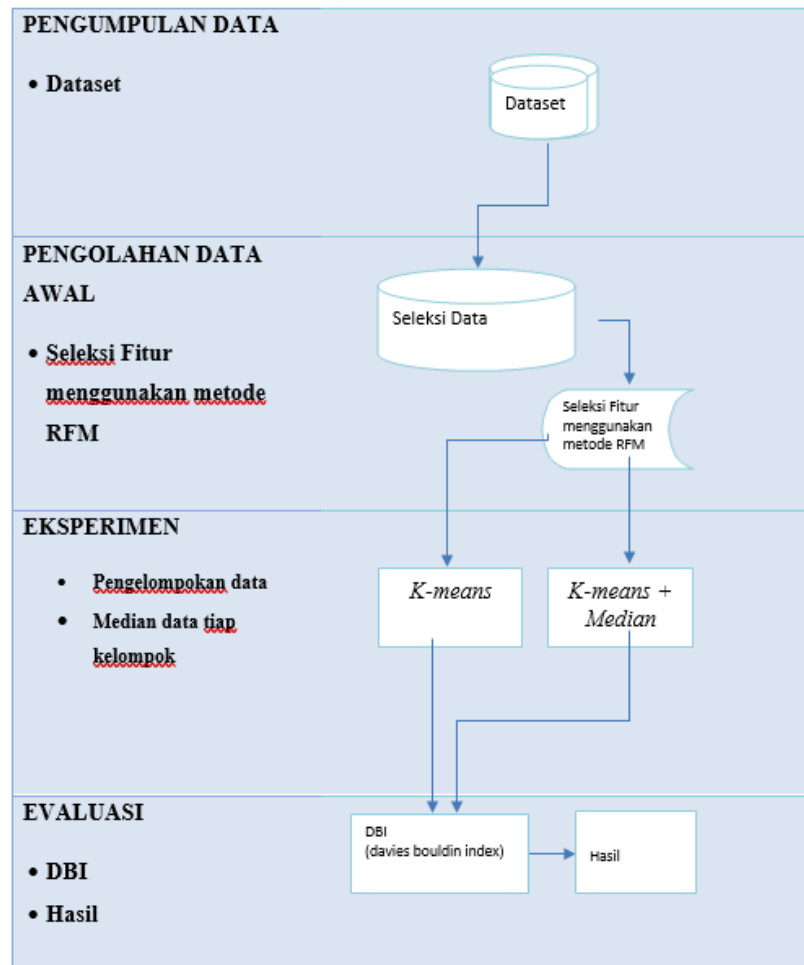
Pada penelitian selanjutnya yang berjudul Pengelompokan Ekuitas Pelanggan Berbasis Recency Frequency Monetary (Rfm) Menggunakan K-Means Clustering [15]. Penelitian ini melakukan segmentasi pelanggan berdasarkan analisis RFM, yaitu recency, frequency dan monetary. Segmentasi loyalitas menggunakan metoda K-Means Clustering. Data transaksi yang digunakan sebanyak 20.814 selama Tahun 2018, dengan jumlah pelanggan 894. Hasil dari penelitian ini didapatkan jumlah objek pada cluster-1 sebanyak 7 pelanggan, sedangkan cluster-2 dan cluster-3 masing- masing 368 dan 390. Tingkat akurasi cluster diukur menggunakan Silhouette Coefficient dengan hasil mendekati 1 yang berarti klusterisasi cukup baik. Interpretasi dari RFM menunjukkan 0,92% pelanggan memiliki potensi loyalitas cukup tinggi, sedangkan 48,10% pelanggan memiliki potensi sebagai pelanggan loyal, dan sisanya 50,98 % tingkat loyalitas diragukan.

5. KERANGKA PEMIKIRAN

Tabel 1 Kerangka Pemikiran

Problem Berdasarkan latar belakang permasalahan yang ada di dapat dirumuskan permasalahannya sebagai berikut : Bagaimana mengoptimasi K-means clatering dengan medians untuk mengelompokkan data kualitas pelanggan.		
Sumber Data Data Transaksi penjualan Grosir UD. Budi Luhur Kota Probolinggo		
Seleksi Fitur Model RFM (<i>Recency Frequency Monetary</i>)	METODE	Clasterring Algoritma K-meansberbasis medians
Evaluasi DBI (davies bouldin index)		

Data uji yang dilakukan pada penelitian ini bersumber pada sebuah *grosir* UD. Budi Luhur yang beradi di kota Probolinggo. Data set menggunakan data transaksi internal yang di peroleh dan di ambil pada akhir tahun 2015. Yang terdiri dari 8 fitur antara lain dengan atribut NoJual, TglJual, KodCus, KdBrg, Jumlah, HrgSatuan, Diskon, dan TotalBeli.



Gambar 1 Alur Tahapan Metode yang Diusulkan

6. Hasil Eksperimen

K-Means + Median Dari experimen yang dilakukan terjadi iterasi sebanyak 10 kali untuk menghasilkan cluster terbaik

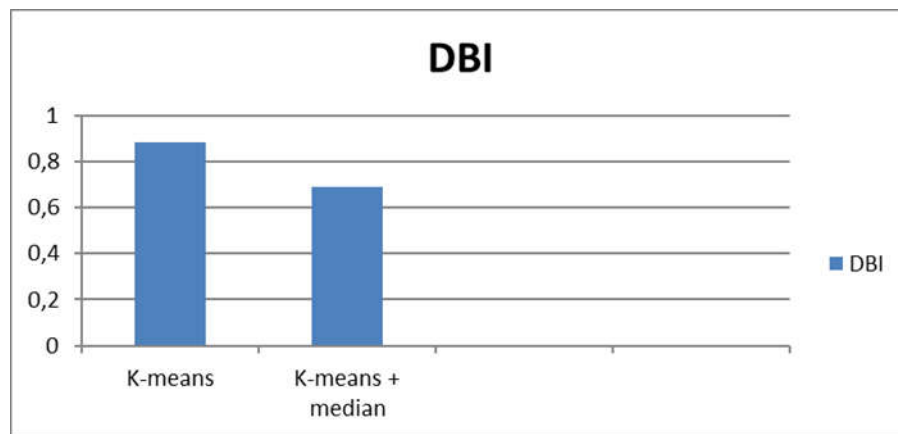
Tabel 2 Hasil K-Means + Median

Jumlah Setiap Cluster	
TOT C1	30
TOT C2	128
TOT C3	136
TOT C4	206

K-Means Dari experimen yang dilakukan terjadi iterasi sebanyak 9 kali untuk menghasilkan cluster terbaik

Tabel 3 Hasil K-Means

Jumlah Setiap Cluster	
TOT C1	128
TOT C2	24
TOT C3	61
TOT C4	287



Gambar 2 Grafik Perbandingan Hasil DBI Antara Metode K-means dan Median K-means

7. KESIMPULAN

Dalam penelitian ini dilakukan pengujian model dengan menggunakan metode k-means dengan pemilihan pusat kluster awal berbasis median untuk mengurangi ketergantungan terhadap pemilihan pusat kluster awal. Hasil eksperimen metode k-means dengan pemilihan pusat kluster awal berbasis median. Dari hasil eksperimen di atas dapat menghasilkan kualitas kluster yang lebih baik karena dengan pemilihan pusat kluster awal berbasis median tidak terjadi kluster kosong karena pemilihan pusat kluster awal yang tidak tepat, walaupun dari segi kecepatan proses masih kalah dari K-means standart. Kluster yang dihasilkan yaitu pada cluster pertama 133 cluster Kedua 30 dan cluster ketiga 208 sedangkan cluster keempat 129 dengan nilai DBI 0,694. Sedangkan metode k-means dengan pemilihan pusat kluster standart didalam experimen ini menggunakan tool rapidminer didapat cluster pertama 310 cluster Kedua 19 dan cluster ketiga 37 sedangkan cluster keempat 134 dengan nilai DBI 0,807.

8. DAFTAR PUSTAKA

[1] Dian Melisa, "Hubungan Antara Citra Merek Dengan Loyalitas Konsumen Produk The

- Body Shop Pada Mahasiswi Universitas Diponegoro,”*Jurnal Empati*, Volume 5(4), 630-633, Oktober 2016.
- [2] Zainuddin Tahuman, “Analisis Faktor-Faktor Yang Mempengaruhi Loyalitas Pelanggan Serta Dampaknya Terhadap Keunggulan Bersaing,”*Jurnal Riset Bisnis dan Manajemen* Vol 4 ,No.3,2016.
- [3] Sudriyanto, “Clustering Loyalitas Pelanggan Dengan Metode RFM (Recency,Frequency,Monetary) dan Fuzzy C-Means,” *Pros. SNATIF Ke-4*, pp. 815–822, 2017.
- [4] Tikaridha Hardiani, Selo Sulisty, Rudy Hartanto, “Segmentasi Nasabah Tabungan Menggunakan Model RFM (Recency, Frequency,Monetary) dan K-Means Pada Lembaga Keuangan Mikro”,*Seminar Nasional Teknologi Informasi dan Komunikasi Terapan (SEMANTIK)*,2015
- [5] E. H. Witten, Ian H. Frank, "Data Mining Practical Machine Learning Tools and Techniques Third Edition". 2011.
- [6] P. Berkhin, “Survey of Clustering Data Mining Techniques,” pp. 1–56.
- [7] E. H. Witten, Ian H. Frank, "Data Mining Practical Machine Learning Tools and Techniques Third Edition". 2011.
- [8] X. Wu and V. K.umar, *The Top Ten Algorithms in Data Mining*. 2009.
- [9] P. J. Flynn, “Data Clustering : A Review,” vol. 31, no. 3, 2000.
- [10] A. K. Jain, “Data clustering : 50 years beyond K-means q,” *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, 2010.
- [11] J. A. Lozano and P. Larra, “An empirical comparison of four initialization methods for the K - Means algorithm,” vol. 20, 1999. [8] M. E. Celebi, H. A. Kingravi, and P. A. Vela, “Expert Systems with Applications A
- [12] M. E. Celebi, H. A. Kingravi, and P. A. Vela, “Expert Systems with Applications A.
- [13] S. Deelers and S. Auwatanamongkol, “Enhancing K-Means Algorithm with Initial Cluster Centers Derived from Data Partitioning along the Data Axis with the Highest Variance,” vol. 26, no. December, pp. 323–328, 2007.
- [14] M. E. Celebi, “A Comparative Study of Efficient Initialization Methods for the K-Means Clustering Algorithm,” vol. 2, no. 1, 2012.
- [15] Yulison Herry Chrisnanto¹, Ade Kanianingsih²,”*Pengelompokan Ekuitas Pelanggan Berbasis Recency Frequency Monetary (Rfm) Menggunakan K-Means Clustering*”,ISSN: 2089-9815,*Seminar Nasional Teknologi Informasi dan Komunikasi*,2019
- [16] Hendi Shopian, “Pemilihan Pusat Klaster Awal Pada Metode K-Means Berbasis Median”,ISSN Cetak: 2087-4286,ISSN On Line: 2580-6017.
- [17] Soumi Ghosh,Sanjay Kumar Dubey,”*Comparative Analysis of K-Means and Fuzzy CMeans Algorithms*”,Vol. 4, No.4, 2013,
- [18] Navjot Kaur, Jaspreet Kaur Sahiwal, Navneet Kaur,”*Efficient K-Means Clustering Algorithm Using Ranking Method In Data Mining*”,ISSN: 2278 – 1323,2012.
- [19] Yusuf Priyo Anggodo, Winda Cahyaningrum, Aprilia Nur Fauziyah, Irma Lailatul Khoiriyah, Oktavianis Kartikasari, Imam Cholissodin, “Hybrid K-Means Dan Particle Swarm Optimization Untuk *Clustering* Nasabah Kredit”, Vol. 4, No. 2, Juni 2017, hlm. 104-110
- [20] Mardalius,”*Implementasi Algoritma K-Means Clustering Untuk Menentukan Kelas Kelompok Bimbingan Belajar Tambahan*”, *Proceeding SEMILOKA ROYAL 2017 “Teknologi Mobile*”,2017.
- [21] Hurriyati, Ratih. 2005. *Bauran Pemasaran dan Loyalitas Konsumen*. Edisi 1. Bandung: Alfabeta