



Comparison K-Means K-Medoids and Fuzzy C-Means for Clustering Customer Data with LRFM Model

Perbandingan K-Means K-Medoids dan Fuzzy C-Means untuk Pengelompokan Data Pelanggan dengan Model LRFM

**Hamdi Syukron^{1*}, Muhammad Fauzi Fayyad², Farin Junita Fauzan³,
Yulia Ikhsani⁴, Umairah Rizky Gurning⁵**

^{1,2,3,4,5} Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,
UIN Sultan Syarif Kasim Riau, Indonesia

^{1,3,5} Puzzle Research Data Technology, Fakultas Sains dan Teknologi,
UIN Sultan Syarif Kasim Riau, Indonesia

E-Mail: ¹hamdisyukron650@gmail.com, ²fauzifayyad@gmail.com, ³12050323918@students.uin-suska.ac.id, ⁴12050320529@students.uin-suska.ac.id, ⁵11950320687@students.uin-suska.ac.id

Received Sept 02nd 2022; Revised Sept 10th 2022; Accepted Sept 20th 2022
Corresponding Author: Hamdi Syukron

Abstract

Indonesia has a potential market for cosmetic companies because it has a population of nearly 270 million people. The growth of the cosmetic industry in Indonesia has experienced rapid development with a percentage growth of 5.59% in August 2021. With this growth, cosmetic companies have resellers spread throughout Indonesia. This study uses customer data from a beauty company reseller. Common problems that often occur in selling beauty products include which customers shop often, which products are often purchased by customers, and which customers are the most loyal. Length, Recency, Frequency, and Monetary or LRFM is a method used to measure the value of a customer based on the history of transactions that have been made. This research was conducted to group customer data based on time span and number of purchase transactions using a clustering algorithm, namely, k-means, k-medoids, and fuzzy c-means. It is hoped that this research can determine the best algorithm in data grouping by comparing the K-Means, K-Medoids, and Fuzzy C-Means clustering algorithms with the LRFM model. The results showed that in clustering customer data using the LRFM model, the K-Means algorithm was superior to K-Medoids and Fuzzy C-means in clustering customer data, by proving the best DBI validity value with a score of 0.167 for the number of clusters, which is 6.

Keyword: Comparison, Fuzzy C-Means, K-Means, K-Medoids, LRFM

Abstrak

Indonesia memiliki pasar yang potensial untuk perusahaan kosmetik karena memiliki jumlah penduduk yang berjumlah hampir 270 juta jiwa. Pertumbuhan industri kosmetik di Indonesia mengalami perkembangan yang pesat dengan persentase pertumbuhan 5,59% pada bulan agustus 2021 silam. Dengan pertumbuhan tersebut perusahaan kosmetik memiliki reseller yang tersebar diseluruh daerah Indonesia. Penelitian ini menggunakan data pelanggan dari salah satu reseller perusahaan kecantikan. Pelanggan mana yang sering berbelanja, produk mana yang sering mereka beli, dan klien mana yang paling setia adalah masalah umum yang sering muncul saat menjual produk kecantikan. Panjang, Kekinian, Frekuensi, dan Moneter, kadang-kadang dikenal sebagai LRFM, adalah teknik yang digunakan untuk menghitung nilai pelanggan berdasarkan riwayat transaksi mereka. Penelitian ini dilakukan untuk mengelompokan data pelanggan berdasarkan rentang waktu dan jumlah transaksi pembelian menggunakan algoritma clustering yaitu, k-means, k-medoids, dan fuzzy c-means. Harapannya, penelitian ini dapat ditentukan algoritma terbaik dalam pengelompokan data dengan membandingkan algoritma clustering K-Means, K-Medoids, dan Fuzzy C-Means dengan model LRFM. Hasil penelitian menunjukkan bahwa dalam pengelompokan data pelanggan dengan model LRFM algoritma K-Means lebih unggul dibanding K-Medoids dan Fuzzy C-means pada mengklasterisasi data pelanggan, dengan dibuktikan pada nilai validitas DBI terbaik dengan perolehan nilai yaitu 0,167 pada jumlah klaster yaitu 6.

Kata Kunci: Fuzzy C-Means, K-Means, K-Medoids, Perbandingan.

1. PENDAHULUAN

Berdasarkan Data dari Badan Pusat Statistik, populasi Negara Indonesia berjumlah hampir 270 juta jiwa [1] menjadikan Indonesia pasar yang potensial untuk perusahaan kosmetik [2]. Setiap tahunnya, pertumbuhan industri kosmetik di Indonesia mengalami perkembangan yang pesat [3]. Data perkembangan pasar pada Bulan agustus 2021 silam menunjukkan industri kosmetik tumbuh sebesar 5,59%. Karena perkembangan tersebut, banyak perusahaan produk kecantikan yang memiliki *reseller* yang tersebar diseluruh daerah di Indonesia. *Reseller* ialah proses bisnis dengan sistem penjualan kembali produk suatu brand [4][5]. Pengecer biasanya memperhatikan masalah yang sering muncul dalam penjualan mereka, seperti konsumen mana yang sering berbelanja, produk mana yang sering mereka beli, dan klien mana yang paling setia.

Length, *Recency*, *Frequency*, dan *Monetary* atau yang disebut dengan LRFM digunakan untuk menghitung nilai pelanggan berdasarkan riwayat transaksi mereka. Panjang antara transaksi awal dan akhir adalah jarak yang ada dari waktu ke waktu. Periode antara transaksi pertama dan terakhir pelanggan dikenal sebagai kebaruan. Jumlah total transaksi pelanggan selama periode waktu tertentu dikenal sebagai frekuensi. Finansial adalah efek akhir dari penggunaan uang konsumen selama periode waktu tertentu.[6]. Pada penelitian ini menggunakan data pelanggan pada salah satu *reseller* dari perusahaan kecantikan berdasarkan rentang waktu dan jumlah transaksi pembelian. Penelitian ini dilakukan klusterisasi pada data pelanggan tersebut menggunakan algoritma *clustering* yaitu, k-means, k-medoids, dan fuzzy c-means dan membandingkannya untuk mencari algoritma terbaik dalam pengelompokan data pelanggan.

Clustering merupakan metode analisa data dengan mengelompokkan data berdasarkan kecocokan karakteristiknya [7]. Algoritma *clustering* yang biasa digunakan adalah K-Means, K-Medoids, dan Fuzzy C-Means. Algoritma K-Means membagi sebuah objek menjadi k-cluster berdasarkan seberapa kompatibel atau tidak kompatibelnya objek tersebut, dan mengelompokkan posisi objek berdasarkan nilai mean terdekat. Sebuah pendekatan yang disebut K-Medoids, yang didasarkan pada penggunaan medoid, mencoba untuk mengurangi sensitivitas partisi sehubungan dengan nilai-nilai ekstrim dalam dataset. Dalam metode pengelompokan yang dikenal sebagai fuzzy C-Means, tingkat partisipasi dalam setiap kelompok mempengaruhi apakah suatu titik data ada [8].

Rizky Gurning dan Mustakim, 2021 sebelumnya melakukan penelitian mengenai penerapan algoritma K-Means dan K-Medoids untuk pengelompokan data Covid-19 dengan atribut jenis penularan, usia, jenis kelamin, paskes dan kecamatan. Analisis hasil pada penelitian ini didapatkan bahwa algoritma terbaik untuk pengelompokan data Covid-19 adalah k-means dengan perolehan nilai terbaik validasi DBI yaitu sebesar 0,139 pada jumlah cluster 4 [9]. Berikutnya, penelitian oleh Ahmad Jaini dkk, 2020 dimana pada penelitian ini membandingkan dua algoritma yaitu fuzzy c-means dan k-medoids untuk mencari algoritma terbaik dalam pengelompokan data penjualan pada 212 mart. Hasil daripada penelitian yang dilakukan oleh Jaini yaitu algoritma fuzzy c-means adalah algoritma terbaik dalam pengelompokan data dengan perolehan nilai SI yaitu dengan perolehan nilai 0,2159 sementara algoritma k-medoids memperoleh nilai SI yaitu sebesar 0,2018 [7]. Selanjutnya adalah penelitian yang dilakukan oleh Ni Putu Viona dkk, 2022 dimana ia melakukan penelitian mengenai pengelompokan data menggunakan algoritma k-means dengan model LRFM untuk pemetaan data pelanggan sebagai syarat memenuhi strategi pengelolaan. Pada penelitian ini model *Length*, *Recency*, *Frequency*, dan *Monetary* atau LRFM digunakan sebagai pemetaan data dengan menyesuaikan atribut dari LRFM. Penelitian ini memiliki hasil analisis *cluster* terbaik menggunakan algoritma k-means dengan perolehan nilai DBI 0,135 pada jumlah cluster yaitu 5 [10].

Oleh karena itu, penulis mengangkat tema berdasarkan latar belakang yang telah dipaparkan sebelumnya yakni “Perbandingan K-Means, K-Medoids, dan Fuzzy C-Means untuk Pengelompokan Data Pelanggan dengan Model LRFM.”

2. BAHAN DAN METODE

Penelitian ini dilalui dengan lima tahapan untuk mendapatkan hasil dan analisis dimulai dari pengumpulan data, pre-processing data setelah itu dilakukan pemodelan kedalam bentuk *Length*, *Recency*, *Frequency* and *Monetary* atau LRFM serta tahapan proses klusterisasi dan juga validitas *cluster* pada algoritma K-Means, K-Medoids, dan Fuzzy C-Means.

2.1 K-Means

Algoritma K-Means ialah salah satu algoritma clustering dalam data mining untuk melakukan pengelompokan data. Algoritma K-Means dapat mempartisi data menjadi dua kelompok ataupun lebih berdasarkan pengelompokan data nonhierarki (sekatan). Metode ini akan mengelompokkan data ke dalam kelompok yang memiliki karakteristik data yang sama sedangkan data dengan karakteristik yang berbeda akan ditambahkan ke kelompok lainnya [11][12].

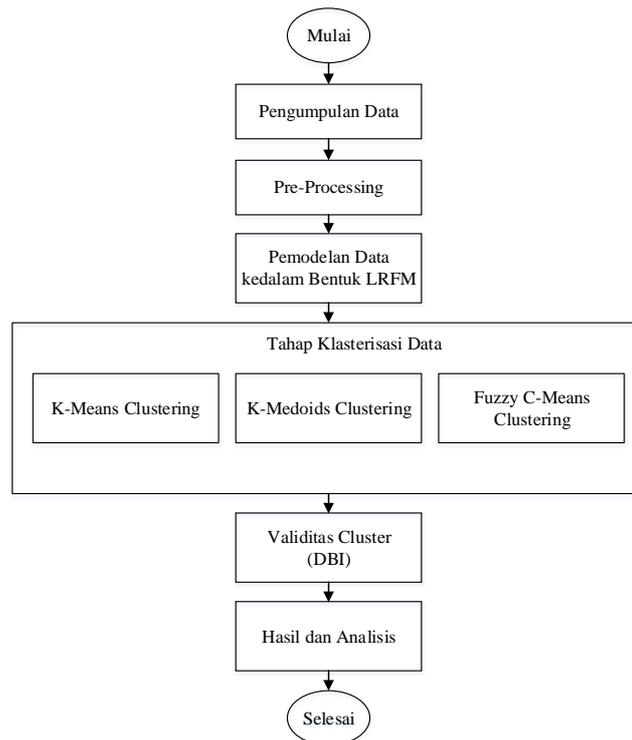
$$d(X_{ij}, C_{kj}) = \sqrt{\sum_{j=1}^m (X_{ij} - C_{kj})^2} \quad (1)$$

K-Means akan mengelompokkan elemen data kedalam suatu kluster dengan jarak terdekat dengan melakukan uji tiap elemen data pada sekumpulan data yang dimiliki. Untuk menghitung hal ini dapat menerapkan persamaan 2 [12].

$$m \sum_k^k d (X_{ij}, C_{kj}) = \sqrt{\sum_{j=1}^m (X_{ij} - C_{kj})^2} \quad (2)$$

Keterangan :

- $d (X_{ij}, C_{kj})$ = Jarak *Euclidian Distance* antara pengamatan ke-i variabel ke-j ke pusat *cluster* ke-k pada variabel ke j.
 x_{ij} = objek pada pengamatan ke-i pada variabel ke-j,
 c_{kj} = pusat kelompok ke-k pada variabel ke-j,
 m = banyaknya variabel yang diamati.



Gambar 1. Tahapan Penelitian

2.3 K-Medoids

Untuk menemukan titik pusat dari suatu cluster maka digunakan sebuah algoritma yaitu k-medoids. K-Medoids hampir sama dengan K-Means perbedaannya K-Medoids ditujukan untuk mengurangi kepekaan dan partisi nilai – nilai ekstrem pada sebuah dataset. Setiap data ditempatkan ke cluster terdekat dengan persamaan Euclidian pada persamaan 3 [11][13].

$$d (X_{ij}, C_{kj}) = \sqrt{\sum_{j=1}^m (X_{ij} - C_{kj})^2} \quad (3)$$

Keterangan :

- $d (X_{ij}, C_{kj})$ = Jarak *Euclidian Distance* antara pengamatan ke-i variabel ke-j ke pusat *cluster* ke-k pada variabel ke j.
 x_{ij} = objek pada pengamatan ke-i pada variabel ke-j,
 c_{kj} = pusat kelompok ke-k pada variabel ke-j,
 m = banyaknya variabel yang diamati.

2.4 Fuzzy C-Means (FCM)

Metode Fuzzy C-Means mengelompokkan data berdasarkan derajat keanggotaan [16][17][18]. Data dapat dikelompokkan berdasarkan derajat keanggotaannya, yaitu berkisar antara 0 sampai 1, dan ada beberapa tipe data yang hanya menampilkan keanggotaan sebagian. Fuzzy Clustering digunakan oleh fuzzy C-Means untuk menetapkan kepemilikan data pada setiap cluster yang masing-masing memiliki keanggotaan yang berbeda. Derajat keanggotaan mengontrol rentang antara 0 dan 1 keberadaan data dalam cluster. Algoritma

Fuzzy C-Means memiliki keuntungan yang sangat baik dalam mendeteksi cluster tingkat tinggi dan mengungkapkan hubungan antara berbagai model cluster. Fungsi fitness atau fungsi objektif dapat diminimalkan agar mendapatkan struktur kluster dan juga menghasilkan nilai kluster yang optimal. Fungsi objektif dilakukan pada persamaan seperti berikut[8][14].

$$jm = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^m d^2(y_{1,z1}) \tag{4}$$

$$u_{ij} = \left[\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad 1 \leq i \leq c ; 1 \leq j \leq n \tag{5}$$

$$z_i = \frac{\sum_{j=1}^n (u_{ij})^m y_j}{\sum_{j=1}^n (u_{ij})^m} \tag{6}$$

2.5. Davies-Bouldin Index (DBI)

DBI bertujuan untuk meminimalkan jarak antar anggota cluster sekaligus memaksimalkan jarak antar anggota cluster [9] [15]. Pengukuran DBI adalah nilai rata-rata yang dapat dibandingkan antara setiap cluster dan yang paling sebanding. Strategi *clustering* terbaik ditunjukkan dengan nilai DBI yang berkurang atau nilai positif terkecil yang mendekati 0. (refumai). Persamaan 7 dapat digunakan untuk menentukan nilai DBI [9].

$$DBI = \sum_{i=1}^p \frac{\sigma_i + \sigma_i}{p} \tag{7}$$

2.6. Length, Recency, Frequency, and Monetary (LRFM)

Pada tahun 1994 Hughes memperkenalkan model RFM yaitu model berbasis perilaku dan dikembangkan kemudian Chang dan Tsay memberikan usul untuk mengambil panjang relasi untuk memecahkan masalah pada data pelanggan sehingga menjadi metode LRFM. Mengetahui lama transaksi nasabah selama periode waktu tertentu (L/Length) penting untuk menghitung loyalitas pelanggan. Fitur model LRFM meliputi *Transaction Length* (L) yang menunjukkan jarak dari waktu ke waktu antara transaksi awal pelanggan dan transaksi terakhir (hari) *Recent Transaction* (R) yang mengukur interval antara transaksi terakhir yang dilakukan konsumen dan waktu analisis saat pengguna melakukan segmentasi (hari), *Annual Frequency* (F) yang menunjukkan volume khas transaksi yang dilakukan oleh klien selama periode waktu tertentu, *Average Monetary Value* (M) yaitu jumlah rata-rata uang yang dipertukarkan oleh konsumen selama periode tertentu [15][16][17].

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan dan Preprocessing Data

Penelitian ini mengambil data pelanggan dari seorang *reseller* di Provinsi Riau berdasarkan rentang waktu pembelian dan jumlah transaksi yang akan menjadi objek penelitian ini. Data preprocessing dilakukan dengan 3 tahap yakni cleaning, transformasi dan normalisasi data-data. Tabel 1 ialah hasil dari preprocessing data.

Tabel 1. Data Preprocessing dengan Model LRFM

No.	Lenght	Nilai Recency	Frequency	Monetary
1	0,94	1,00	0,88	1,00
2	0,96	0,71	0,25	0,08
3	0,98	0,00	0,00	0,11
4	0,11	0,25	0,13	0,05
5	0,00	0,69	0,38	0,49
6	0,63	0,07	0,13	0,30
7	0,00	0,62	0,25	0,13
8	0,00	0,51	0,00	0,01
...
29	0,50	0,95	1,00	0,57
30	0,33	0,89	0,38	0,58

3.2 Proses Clustering dan Validitas Cluster

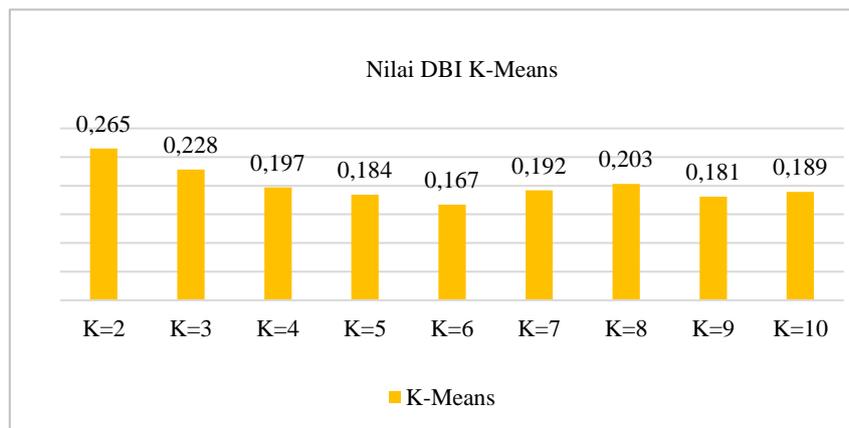
Tiga algoritma yang digunakan untuk proses clustering yaitu K-Means, K-Medoids dan Fuzzy C-Means. Dari 30 data pelanggan hasil processing sebelumnya dilakukan pengolahan data dengan tiga algoritma tersebut.

3.2.1. K-Means

Tabel 2. Hasil Klasterisasi K-Means

Percobaan	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10
K = 2	8	22	-	-	-	-	-	-	-	-
K = 3	11	6	13	-	-	-	-	-	-	-
K = 4	5	13	8	4	-	-	-	-	-	-
K = 5	7	4	6	9	4	-	-	-	-	-
K = 6	4	1	9	3	6	7	-	-	-	-
K = 7	3	7	3	4	4	8	1	-	-	-
K = 8	9	1	3	2	2	4	7	2	-	-
K = 9	3	2	6	9	2	4	1	2	1	-
K = 10	5	2	8	3	2	4	1	2	2	1

Setelah proses klasterisasi dengan algoritma k-means selanjutnya tahap validitas cluster menggunakan metode DBI. gambar 2 adalah hasil klasterisasi data menggunakan algoritma K-Means dengan indeks Nilai DBI.



Gambar 2. Nilai DBI K-Means

Eksperimen dengan k=6 dan nilai DBI 0,167 menghasilkan cluster terbaik pada algoritma k-means, yaitu membagi data menjadi 6 cluster dengan 4 anggota di cluster 1, 1 anggota di cluster 2, 9 anggota di cluster 3, 3 anggota di cluster 4, dan 7 anggota di cluster 6.

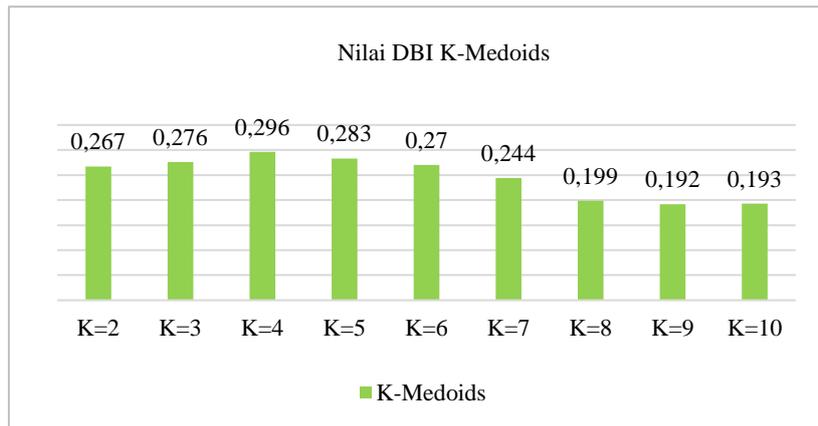
3.2.2. K - Medoids

Selanjutnya penggunaan algoritma K-Medoids untuk proses clustering dengan percobaan *cluster* yang sama dengan algoritma sebelumnya, sehingga juga dilakukan eksperimen *cluster* 2, 3, 4, 5, 6, 7, 8, 9 dan 10. Tabel 3 ialah hasil klasterisasi dengan algoritma K-Medoids.

Tabel 3. Hasil Klasterisasi K-Medoids

Percobaan	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10
K = 2	18	12	-	-	-	-	-	-	-	-
K = 3	10	9	11	-	-	-	-	-	-	-
K = 4	6	9	9	6	-	-	-	-	-	-
K = 5	6	6	9	4	5	-	-	-	-	-
K = 6	4	5	9	4	3	5	-	-	-	-
K = 7	4	5	9	4	4	2	2	-	-	-
K = 8	1	5	9	4	3	4	2	2	-	-
K = 9	5	1	2	5	3	1	9	1	3	-
K = 10	5	1	2	5	3	1	2	7	1	3

Kemudian dilakukan tahap validitas cluster dengan metode Davies Bouldin Index (DBI). Gambar 3 adalah hasil klasterisasi data menggunakan algoritma K-Medoids dengan indeks Nilai DBI.



Gambar 3. Nilai DBI K-Medoids

Berdasarkan indeks nilai DBI, eksperimen dengan k=9 dan nilai DBI 0,192 memiliki cluster terbaik menggunakan algoritma k-medoids, diikuti oleh eksperimen dengan k=10 dan nilai DBI 0,193. Nilai eksperimen k=9 memisahkan data menjadi sembilan cluster, dengan lima individu di cluster 1, satu di cluster 2, dua di cluster 3, lima di cluster 4, tiga di cluster 5, dan empat di cluster 6, sembilan di cluster 7, satu di cluster 8, dan satu di cluster 9.

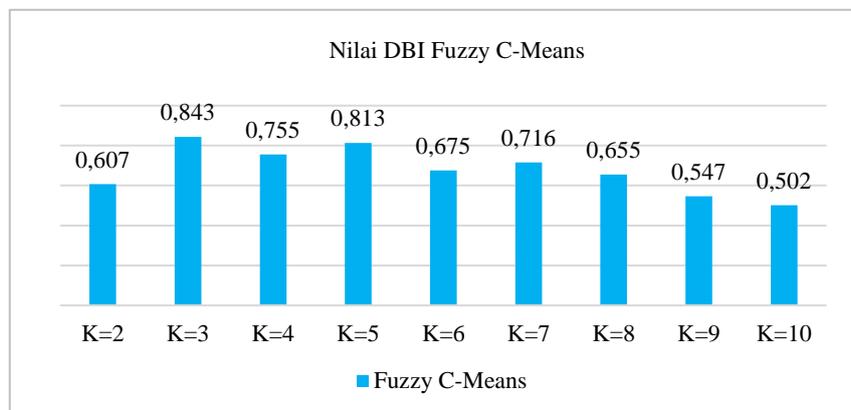
3.2.3. Fuzzy C-Means

Terakhir penggunaan algoritma Fuzzy C-Means untuk proses clustering dengan eksperimen yang sama dengan algoritma sebelumnya, sehingga juga menggunakan percobaan cluster k=2 hingga k=10. Tabel 4 ialah hasil klasterisasi dengan algoritma Fuzzy C-Means.

Tabel 4. Hasil Klasterisasi Fuzzy C-Means

Percobaan	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10
K = 2	11	19	-	-	-	-	-	-	-	-
K = 3	16	10	4	-	-	-	-	-	-	-
K = 4	10	3	7	10	-	-	-	-	-	-
K = 5	11	5	6	3	5	-	-	-	-	-
K = 6	8	6	4	3	5	4	-	-	-	-
K = 7	3	3	4	3	5	4	8	-	-	-
K = 8	4	5	5	4	3	3	2	4	-	-
K = 9	3	3	3	3	6	3	3	4	2	-
K = 10	3	3	2	1	3	4	8	2	2	2

Kemudian tahap validitas cluster menggunakan metode DBI. Gambar 4 adalah hasil klasterisasi data menggunakan algoritma Fuzzy C-Means dengan indeks Nilai DBI.



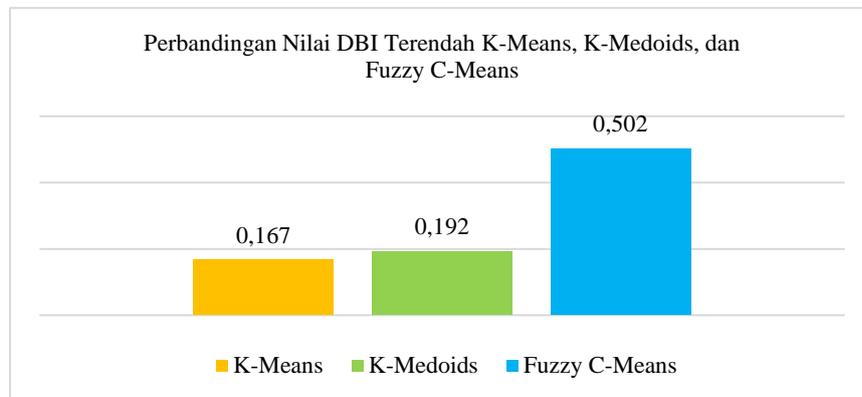
Gambar 4. Nilai DBI Fuzzy C-Means

Percobaan dengan k=10 dan nilai DBI 0,502 memiliki cluster terbaik menurut metode fuzzy c-means, menurut indeks nilai DBI. Dengan k=10, eksperimen membagi data menjadi 10 cluster, masing-masing dengan

tiga anggota (cluster 1), dua anggota (cluster 2), satu anggota (cluster 3), tiga anggota (cluster 4), empat anggota (cluster 5), delapan anggota (cluster 7), dua anggota (cluster 8), dua anggota (cluster 9), dan dua anggota (cluster 10).

3.3. Perbandingan Hasil DBI Algoritma K-Means, K-Medoids, dan Fuzzy C-Means

Gambar 5 ialah Grafik Perbandingan indeks nilai DBI tertinggi untuk setiap algoritma, algoritma K-Means dengan indeks nilai DBI sebesar 0,167 pada $k=6$, algoritma k-medoids dengan nilai DBI sebesar 0,192 pada $k=9$ dan algoritma Fuzzy C-Means dengan nilai DBI sebesar 0,502 pada $k=10$.



Gambar 5. Perbandingan Nilai DBI Terendah K-Means, K-Medoids dan Fuzzy C-Means

Gambar 5 menunjukkan Algoritma K-Means mempunyai nilai validitas cluster terbaik dibandingkan algoritma K-Medoids dan Fuzzy C-Means. sehingga algoritma cluster paling optimal adalah K-Means dengan nilai $k=6$.

4. KESIMPULAN

Hasil analisis dan pengolahan data dengan model *Length*, *Recency*, *Frequency*, dan *Monetary* atau LRFM untuk perbandingan algoritma clustering k-means, k-medoids, dan fuzzy c-means menunjukkan bahwasannya *cluster* terbaik untuk pengelompokan data pelanggan ialah pada algoritma k-means pada jumlah kluster yaitu 6. Sementara itu, algoritma k-medoids dan fuzzy c-means perolehan nilai *cluster* terbaik yaitu pada jumlah kluster 9 dan 10. Selanjutnya, dilakukan validasi menggunakan nilai indeks Davies Bouldin Index (DBI) berdasarkan hasil klusterisasi tersebut. Maka pada penelitian ini didapatkan algoritma terbaik dalam pengelompokan data pelanggan dengan model LRFM yaitu k-means lebih unggul daripada k-medoids dan fuzzy c-means dengan nilai indeks DBI terbaik sebesar 0,167 pada jumlah kluster 6. Eksperimen dengan $k=6$ dan nilai DBI 0,167 menghasilkan cluster terbaik pada algoritma k-means, yaitu membagi data menjadi 6 cluster dengan 4 anggota di cluster 1, 1 anggota di cluster 2, 9 anggota di cluster 3, 3 anggota di cluster 4, dan 7 anggota di cluster 6. Jadi, pola data yang didapatkan pada penelitian ini dapat dijadikan sebagai acuan gambaran model clustering untuk data pelanggan.

UCAPAN TERIMAKASIH

Ucapan terimakasih kepada salah satu *reseller* di Riau yang telah memberikan kesempatan kepada kami untuk pengambilan data yang bertujuan sebagai bahan penelitian. Terimakasih juga kepada Puzzle Research Data and Technology (Predatech) yang telah memberikan masukan, dorongan dan semangat dalam penyelesaian penelitian ini.

REFERENSI

- [1] "Badan Pusat Statistik." <https://sulut.bps.go.id/indicator/12/958/1/jumlah-penduduk-menurut-provinsi-di-indonesia.html> (accessed Jun. 18, 2022).
- [2] "Analisis Strategi Pemasaran Produk Kosmetik Wardah dengan Pendekatan SWOT-AHP (ANALYTIC HIERARCHY PROCESS)," *JURNAL SAINS DAN SENI ITS*, vol. 6, p. 1, 2017.
- [3] D. Wulan Amalia and D. Amanda Zuliestiana, "PENGARUH ELECTRONIC WORD OF MOUTH DI MEDIA SOSIAL YOUTUBE TERHADAP MINAT BELI PRODUK KOSMETIK NEW WARDAH EXCLUSIVE SERIES (STUDI KASUS KOTA-KOTA DI JAWA BARAT) *Jurnal Mitra Manajemen (JMM Online)*," *Destiana Wulan Amalia*, vol. 1, no. 6, pp. 871–884, 2020.
- [4] P. Ilmu Komunikasi and J. Ilmu Sosial, "PENGARUH AKTIVASI MEREK MELALUI INSTAGRAM @MSGLOWBEAUTY TERHADAP KETERLIBATAN RESELLER PADA PRODUK MS.GLOW Deby Nur Winda Sari," 2021.

- [5] “Pengaruh Giveaway, Price Discount, dan Reseller terhadap Peningkatan Penjualan Ms.”.
- [6] M. I. Chanafi, D. P. Hapsari, R. K. Hapsari, and T. Indriyani, “Implementasi Algoritma Clustering Untuk Pengelompokan Pelanggan Retail Berdasarkan Skor Recency, Frequency, Dan Monetary.”
- [7] A. Jaini, Mustakim, A. Weni Syaputri, T. Qurahman, S. Thaufik Rizaldi, P. H. Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau Jl Soebrantas Km, and P. Pekanbaru Riau, “Perbandingan Algoritma Fuzzy C-Means dan K-Medoids untuk Pengelompokan Data Penjualan pada 212 Mart,” 2020.
- [8] Sarah S and Mustakim, “Analisis Penerimaan Vaksin Covid-19 Berbasis Fuzzy Clustering Machine Learning di Provinsi Riau,” *Jurnal Riset Komputer*, vol. 8, no. 6, pp. 2407–389, 2021, doi: 10.30865/jurikom.v8i6.3636.
- [9] U. Rizky Gurning and Mustakim, “Penerapan Algoritma K-Means dan K-Medoid untuk Pengelompokan Data Pasien Covid-19,” *Technology and Science (BITS)*, vol. 3, no. 1, 2021, doi: 10.47065/bits.v3i1.1003.
- [10] Ni Putu Viona Viandari, I Made Agus Dwi Suarjaya, and I Nyoman Piarsa, “Pemetaan Pelanggan dengan LRFM dan Two Stage Clustering untuk Memenuhi Strategi Pengelolaan,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 130–139, Feb. 2022, doi: 10.29207/resti.v6i1.3778.
- [11] E. Triwira Lestari, Mustakim and J. Adhiva, “SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Implementation Naive Bayes Classifier Algorithm and K-Nearest Neighbor For Obesity Nutritional Status of Children with Disabilities Implementasi Algoritma Naive Bayes Classifier dan K-Nearest Neighbor Untuk Klasifikasi Status Gizi Obesitas Anak Disabilitas”, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [12] R. Kesuma Dinata, N. Hasdyna, and N. Azizah, “Analisis K-Means Clustering pada Data Sepeda Motor,” 2020.
- [13] S. Sindi *et al.*, “ANALISIS ALGORITMA K-MEDOIDS CLUSTERING DALAM PENGELOMPOKAN PENYEBARAN COVID-19 DI INDONESIA,” (*JurTI*) *Jurnal Teknologi Informasi*, vol. 4, no. 1, pp. 166–173, Jun. 2020, doi: 10.36294/JURTI.V4I1.1296.
- [14] S. Fitri Octavia and M. Mustakim, “Penerapan K-Means dan Fuzzy C-Means untuk Pengelompokan Data Kasus Covid-19 di Kabupaten Indragiri Hilir,” *Technology and Science (BITS)*, vol. 3, no. 2, 2021, doi: 10.47065/bits.v3i2.1005.
- [15] C. Jamunadevi, S. Tamil Selvan, M. Govindarajan, C. Saravanan, and B. R. Janaki Raman, “LRFM model for customer purchase behaviour using K-Means algorithm,” *IOP Conf Ser Mater Sci Eng*, vol. 1055, no. 1, p. 012111, Feb. 2021, doi: 10.1088/1757-899x/1055/1/012111.
- [16] F. Marisa, S. S. S. Ahmad, Z. I. M. Yusof, Fachrudin, and T. M. A. Aziz, “Segmentation model of customer lifetime value in Small and Medium Enterprise (SMEs) using K-Means Clustering and LRFM model,” *International Journal of Integrated Engineering*, vol. 11, no. 3, pp. 169–180, 2019, doi: 10.30880/ijie.2019.11.03.018.
- [17] S. H. Chao, M. K. Chen, and H. H. Wu, “An LRFM Model to Analyze Outpatient Loyalty From a Medical Center in Taiwan,” *Sage Open*, vol. 11, no. 3, 2021, doi: 10.1177/21582440211031899.