Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN: 2548-964X
Vol. 5, No. 12, Desember 2021, hlm. 5278-5287 http://j-ptiik.ub.ac.id

Analisis Metode Fuzzy C-Means dengan Model LRFM terhadap Segmentasi Pelanggan pada Toko Butik (Studi Kasus: Diva Collection Kota Cilegon)

Adityo Dwirahmawan[1], Nanang Yudi Setiawan[2], Djoko Pramono[3]

Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: [1]adityo.dwirahmawan@gmail.com, [2]nanang@ub.ac.id, [3]djoko.jalin@ub.ac.id

Abstrak

Toko butik Diva Collection adalah toko butik yang menjual berbagai macam jenis pakaian wanita yang berdiri pada tahun 2019 di Kota Cilegon. Toko tersebut masih belum memberlakukan strategi Manajemen Hubungan Pelanggan dengan memberikan pelayanan yang berbeda ke setiap pelanggan, sehingga pihak butik memperlakukan sama kepada tiap pelanggan. oleh karena itu, diperlukan segmentasi pelanggan untuk mengetahui bagaimana sifat atau karakteristik pelanggan dengan kesamaan tertentu. Hal ini berguna untuk mencari informasi terkait pelanggan yang menguntungkan. Perilaku pelanggan yang diamati adalah dari proses LRFM. Length(selisih waktu transaksi terakhir dengan awal), Recency (selisih waktu pembelian dengan penelitian), Frequency (jumlah pembelian), dan Monetary (total pengeluaran konsumen). Metode klasterisasi yang digunakan adalah C-Means untuk melakukan segmentasi pelanggan. hasil grafik K-Means dan Silhoutte digunakan untuk menentukan berapa jumlah cluster yang optimal sebelum klasterisasi. Data yang digunakan adalah data periode Agustus 2019 hingga Desember 2019 sebanyak 560 data transaksi pelanggan. jumlah kelompok yang dihasilkan adalah 2 segmen dan 3 segmen. Kelompok 2 adalah kelompok paling optimal, hal ini berdasarkan pemeringkatan berdasarkan nilai F,M,L tertinggi dan R terendah. Visualisasi Dashboard hasil pengelompokan digambarkan dalam bentuk diagram dan grafik. Terdapat 3 halaman yang menampilkan informasi hasil pengelompokkan dan diberikan kepada pihak Toko Butik Diva Collection untuk dilakukan pengujian. Hasil rata-rata yang didapatkan dari 3 responden adalah 72,5. Nilai tersebut tergolong baik, dimana pihak Butik Diva Collection dapat menerima hasil visualisasi Dashboard tersebut.

Kata kunci: *model LRFM, clustering, Fuzzy C-Means, manajemen hubungan pelanggan, dashboard*

Abstract

Diva Collection boutique shop is a boutique store that sells various types of women's clothing that was established in 2019 in Cilegon City. The store still has not implemented a Customer Relationship Management strategy by providing different services to each customer, so that the boutique treats each customer the same. Therefore, customer segmentation is needed to find out how the nature or characteristics of customers with certain similarities are. This is useful for finding information about profitable customers. Observed customer behavior is from the LRFM process. Length (difference between the last transaction and the beginning), Recency (difference between purchase dan research), Frequency (the amount of purchase), and Monetary (total consumer spending). The clustering method used is C-Means to segment customers. The results of the K-Means and Silhoutte graphs are used to determine the optimal number of clusters before clustering. The data used is data for the period August 2019 to December 2019 as many as 560 customer transaction data. the number of groups produced is 2 segments and 3 segments. Group 2 is the most optimal group, this is based on a ranking based on the highest F, M, L values and the lowest R. The Dashboard visualization of the grouping results is described in the form of diagrams and graphs. There are 3 pages that display information on the results of grouping and are given to the Diva Collection Boutique Store for testing. The average result obtained from 3 respondents is 72.5. This value is classified as good, where the Diva Collection Boutique can receive the results of the Dashboard visualization..

Keywords: *LRFM model, clustering, Fuzzy C-Means, customer relationshop manajement, dashboard*

1. PENDAHULUAN

Butik Diva Collection adalah usaha bisnis *fashion* perorangan yang dimiliki oleh Ibu Widayati. Toko Butik ini terletak di wilayah Kota Cilegon. Tepatnya di Jl. Kawat No. 6. Butik *fashion* ini menjual berbagai macam jenis busana wanita mulai dari kaos, kemeja, *outer*, dan gamis. Sistem penjualannya adalah membuka gerai butik di rumah sendiri dan gerai pada Mall Cilegon dengan penyediaan stok *ready-to-wear* atau menyediakan stok busana dengan jumlah tertentu dan *made-by-request* atau pemesanan tergantung permintaan konsumen. Selain gerai, pemasarannya juga dipromosikan pada WhatsApp. Butik Diva Collection berdiri pada bulan Agustus 2019.

Data yang diambil peneliti diambil dari data pelanggan pada gerai di rumah pemilik butik dan juga pada gerai yang berada di Mall Cilegon. Toko Butik Diva Collection ini belum menerapkan Strategi Bisnis dari Manajemen Hubungan Pelanggan. dalam artian, data hanya digunakan untuk menghitung penghasilan perbulan dan hanya mengetahui berapa jumlah pelanggan yang pada bulan tertentu yang melakukan transaksi. Selain belum memiliki strategi bisnis, Butik Diva Collection juga mengalami kerugian pada bulan bulan tertentu, hal ini dikarenakan penjualan terhadap stok yang tersedia tidak terjual semua atau stok yang tersedia di awal masih tersisa. Berdasarkan masalah tersebut, dilakukan pengelompokkan data untuk mengetahui bagaimana perilaku pelanggan. pengelompokkan data dilakukan dengan menggunakan metode *partitioning method clustering*.

Dalam menerapkan *clustering* metode *Fuzzy C-Means*, variabel LRFM digunakan untuk membantu dalam pemetaan nilai nilai pelanggan yang akan dipakai menjadi acuan dalam penerapan strategi bisnis. LRFM yang dimaksud adalah variabel *Length, Recency, Frequency, dan Monetary.* Implementasi *clustering* diterapkan pada aplikasi R-Studio untuk mendapatkan hasil pemetaan pelanggan sebagai penerapan strategi bisnis. Implementasi pemetaan pelanggan berdasarkan data ID pelanggan sebagai data unik tiap pelanggannya. Visualisasi dashboard adalah tahap yang dibuat setelah mendapatkan hasil *clustering* dari proses *Fuzzy C-Means* pada R-Studio. Pembuatan dashboard dilakukan pada *Google Data Studio.* Hasil dari visualisasi inilah yang diserahkan kepada pihak Butik Diva Collection sebagai rekomendasi dari hasil penelitian.

2. LANDASAN KEPUSTAKAAN

2.1 *Data Mining*

Data Mining merupakan proses statistik atau perhitungan, *machine learning* yang digunakan untuk mengidentifikasi dan ekstraksi informasi yang bermanfaat serta pengetahuan yang terkait dari berbagai penyimpanan besar (Efrain Turban, Jay E. Aronson, 2005). Proses yang biasa dilakuan oleh data mining adalah: deskripsi, prediksi, estimasi, klasifikasi, *clustering*, dan asosiasi(Larose, 2005)

2.2 Manajemen Hubungan Pelanggan

Menurut (Buttle, 2018), perusahaan informasi dan teknologi lebih memilih menggunakan istilah manajemen hubungan pelanggan unutk menggambarkan perangkat lunak yang mengoptimatisasi fungsi dari penjualan dan bisnis layanan MHP adalah suatu cara bagaimana untuk mempertahankan pelanggan yang sudah ada atau memperoleh pelanggan baru, serta mengakuisisi pelanggan agar dapat membantu Butik Diva Collection secara maksimal.

2.3 *Clustering*

Pengelompokan atau yang biasa disebut dengan *clustering*, adlah suatu cara penerapan data mining. cara mengidentifikasi daerah padat dan jarang dalam suatu objek yang bertujuan untuk menemukan pola disrtribusi secara keseluruhan dan korelasi yang menarik antar atribut (Han, 2006). Atau menurut (Larose, 2005), *Clustering* adalah metode mengelompokkan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Sebuah *cluster* adalah kumpulan record yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan data *cluster* lainnya. tujuan *clustering* adalah untuk menghasilkan kelompok objek yang mirip satu sama lain dalam kelompok tertentu.

2.4 *Clustering Fuzzy C-Means (FCM)*

Clustering FCM adlaah metode menemukan kelompok atau *cluster* optimal berdasarkan derajat keanggotaan pada posisi *Euclidean Distance* terhadap *centroid* atau pusat data antar

cluster. *Fuzzy C-Means* akan mengelompokkan data berdasarkan letak titik data dalam suatu kelompok yang ditentukan oleh derajat keanggotaannya. Hasil dari perhitungannya adalah deretan pusat kelompok dan kujmpulan derajat keanggotaan pada setiap data menurut (Rahakbauw, Ilwaru and Hahury, 2017). Cara untuk melakukan proses klasterisasi metode *Fuzzy C-Means* adalag sebagai berikut:

1. Memasukkan data pada cluster X. data berupa matriks a x b, dimana n adalah jumlah sampel data, b adalah atribut tiap data. Xkj adalah data sampel ke k, atribut ke-j.
2. Mencari nilai dari:
   a. Jumlah cluster
   b. Maksimum iterasi
   c. Pangkat pembobot
   d. Error terkecil yg diinginkan
   e. Fungsi objektif awal
   f. Iterasi awal

3. Meningkatkan nilai random sebagai elemen elemen matrix partisi di awal U

$$U_0 = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \cdots & \mu_{1c}(x_c) \\ \vdots & \vdots & & \vdots \\ \mu_{11}(x_1) & \mu_{12}(x_2) & \cdots & \mu_{nc}(x_c) \end{bmatrix}$$

   Matriks di fuzzy c-Means harus mempenuhi persyaratan sebagai berikut

$$\mu_{ik} = [0,1]; \quad (1 \le i \le c; 1 \le k \le n$$

$$\sum_{i=1}^{n} \mu_{ik} = 1; \quad 1 \le i \le c$$

$$0 < \sum_{i=1}^{c} \mu_{ik} < c; \quad 1 \le k \le n$$

   Lalu jumlah tiap atribut dicari:

$$Q_j = \sum_{i=1}^{c} (\mu_{ik})$$

   Lalu mencari nilai

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j}$$

4. Selanjutnya mencari nilai pusat cluster ke-K

$$V_{ij} = \frac{\sum_{k=1}^{n} ((\mu_{ik})^m * X_{kj})}{\sum_{k=1}^{n} (\mu_{ik})^m}$$

$$P_t = \sum_{k=1}^{n} \sum_{i=1}^{c} \left(\left[\sum_{j=1}^{m} (X_{kj} - V_{ij})^2 \right] (\mu_{ik})^m \right)$$

5. Mencari fungsi dari objektif iteraasi ke T menggunakan persamaan berikut:

6. Perubahana matriks partisi dihitung dengan cara :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{p} (X_{kj} - V_{ij})^2 \right]^{\frac{-1}{p-1}}}{\sum_{i=1}^{c} \left[\sum_{j=1}^{p} (X_{kj} - V_{ij})^2 \right]^{\frac{-1}{p-1}}}$$

7. Yang terakhir adalah mencari kondisi untuk berhenti :

   a) Jika ($|Pt - Pt{-}1| < \xi$) atau (t < iterasi maksimal), maka berhenti.
   b) Jika tidak, maka t = t + 1 kemudian ulangi langkah ke-4.

2.5 LRFM (*Length, Recency, Frequency, Monetary*)

LRFM model adalah model RFM yang dikembangkan menggunakan cara yang digunakan untuk melakukan segmentasi pasar berdasarkan analisis nilai nilai tiap pelanggan. Model *RFM* beranggotakan tiga atribut. Yaitu *Recency, Frequency,* dan *Monetary*. dimana *Recency* adalah selisih waktu antara pembelian terakhir pelanggan dengan waktu melakukan penelitian. Frequency adalah intensitas seberapa sering pelanggan tersebut melakukan transaksi. Monetary merupakan jumlh pengeluaran yang dilakukan pelanggan untuk transaksi tersebut dalam bentuk uang. Namun, menurut (Wu, Lin and Liu, 2014), RFM model tidak dapat membedakan pelanggan mana yang memiliki hubungan jangka pendek atau jangka panjang terhadap perusahaan. Oleh karena itu, RFM ini dikembangkan dengan menambahkan *Length* atau selisih waktu antara pembelian terakhir dengan pembelian pertama pelanggan.

2.6 *Elbow Method*

Menurut Mahdulatha 2012, Metode perhitungan Elbow adalah suatu cara yang digunakan untuk menghasilkan informasi untuk menentukan jumlah *cluster* terbagus. Yaitu menggunakan cara melihat persentase hasil perbanadingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik (Mahdulatha, 2012). Hasil persentase yang berbeda dari setiap nilai *cluster* dapat ditunjukkan dengan menggunakan grafik sebagai sumber informasinya. Jika nilai *cluster* pertama dengan nilai *cluster* kedua memberikan nilai penurunan paling besar dan *cluster* kedua ke *cluster* ketiga

menunjukkan penurunan yang signifikan tidak terlalu drastis, maka *cluster* tersebutlah yang terbaik

2.7 *Dashboard*

Dashboard adalah suatu perantara yang menggambarkan/memvisualisasikan hasil dari pekerjaan dalam bentuk grafik atau diagram atau tabel dengan tujuan untuk memudahkan pengguna dalam melihat bisnisnya. Dengan membaca hasil *clustering, dashboard* dapat dijadikan media untuk mengambil keputusan strategi suatu bisnis.

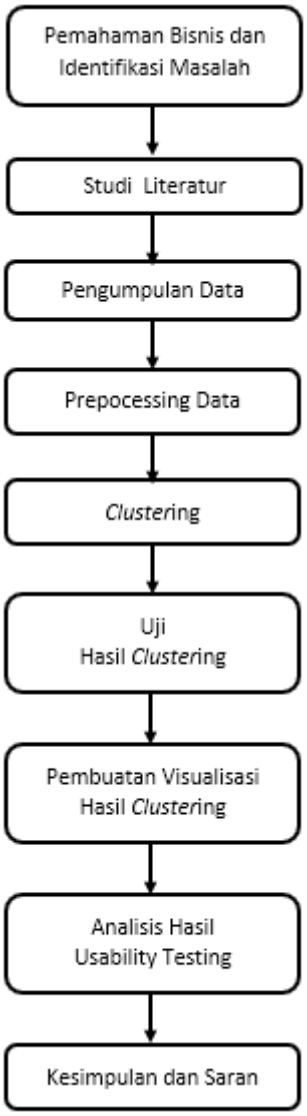
2.8 Usability Testing (System Usability Scale)

Usability Testing metode *System Usability Scale (SUS)* merupakan metode yang paling umum digunakan untuk pengujian terhadap antarmuka atau visualisasi dari sistem yang telah dibuat. Sistem pengujiannya adalah memberikan kuesioner berupa 10 pernyataan kepada responden yang telah dipilih sebelumnya. Dimana pernyataan ganjil adalah pernyataan yang mengandung makna positif dan pernyataan genap merupakan pernyataan yang mengandung makna negatif. Sekala penilaian yang harus diisi oleh responden menggunakan sekala likert rentang 1 sampai 5. Dimana skor 1 menunjukkan penguji sangat tidak meyetujui dan skor 5 menunjukkan kesetujuan dari penguji (Brooke, 2020).

Pernyataan ganjil nilainya didapat dengn cara mengurangi nilai dengann skor 1, pernyataan genap nilainya didapat dengan cara mengurangi skor 5 dengan nilai dari responden. Hasil dari pernyataa genap dan ganjil memiliki nilai rentang 0 sampai 4. Sedangkan hasil dari total pernyataan genap dan ganjil dikalikan dengan nilai 2,5 dan memiliki hasil nilai dengan rentang 0 sampai 100.

3. METODOLOGI

Gambar 1 adalah diagram alur metodologi pada penelitian *Clustering Fuzzy C-Means* Butik Diva Collection.



Gambar 1. Diagram Alur Metodologi

Alur metodologi dimulai dari pemahaman bisnis dan strategi masalah. Hal yang dilakukan adalah mencari informasi berupa data baik dari data transaksi pelanggan ataupun data mengenai masalah yang dialami Butik Diva Collection. Proses yang dilalui melalui wawancara terhadap pihak butik. Setalah itu, dilakukan studi literatur dengan cara mencari referensi dari penelitian yang dapat menyelesaikan masalah serupa dengan masalah butik Diva Collection. Setelah melakukan studi literatur, selanjutnya mengumpulkan data yang diperlukan untuk dijadikan bahan penelitian.

Setelah mengumpulkan data informasi, dilakukan *preprocessing data*. proses yang dilakukan antara lain *cleaning data negatve* dan NA , *transformation data* metode Logaritma, dan Min-Max *normalizazion data*. sebelum memasuki tahap *clustering*, dicari terlebih

dahulu acuan jumlah *cluster* yang optimum. Setelah mendapatkan jumlah kelompok yang optimum, melakukan proses *clustering* dengan metode *Fuzzy C-Means* menggunakan tools R-Studio untuk alat bantu pengolahannya. Hasil dari *clsutering* adalah penyebaran data di beberapa kelompok yang sudah didapatkan.

Pengujian validitas dan verifikasi *cluster* dilakukan setelah mendapatkan hasil *cluster*. Pengujian *validitas* menggunakan metode *Partition Coefficient (PC), Davies Bouldin's Index (DBI), Connectivity, Dunn Index, Silhoutte Index* dan Pengujian SSE. Sedangkan uji verifikasi menggunakan *Euclidean Distance* yang digunakan untuk menghitung jarak data dengan pusat data untuk penempatan data pada *cluster* terdekat. Proses terakhir adalah denormalisasi data untuk menganalisa hasil dari tiap *cluster* untuk mengetahui karakteristik tiap pelanggan pada *clsuter* tersebut.

Hasil analisa tersebut divisualisasikan pada dashboard dalam bentuk diagram batang, lingkaran serta tabel guna mempermudah pihak butik membaca hasilnya. Pemodelan dashboard berdasarkan nilai LRFM dan pemeringkatan *cluster* terbaik hingga terburuk. Selanjutnya adalah penilaian dari pihak butik terhadap visualisasi dashboard menggunakan metode *Usability Testing* jenis *System Usability Scale*. Tahap terakhir adalah pengambilan kesimpulan dan rekomendasi penelitian yang akan datang.

4. Hasil Penelitian

4.1 Ekstraksi Model LRFM

Hasil ekstraksi pemetaan LRFM didapatkan dari proses *recode variable* nilai nilai atribut Length, Recency, Frequency, Monetary dan penghapusan nilai NA serta data negatif pada proses sebelumnya. Nilai variabel *length*, *recency* didapatkan dari variabel ID dan tanggal, variabel *frequency* didapatkan dari variabel ID dan nomor transaksi variabel *monetary* daidapatkan dari variabel ID dan total jumlah uang. Setelah dilakukan ekstraksi, jumlah data perubahan dari data mentah sebanyak 560 data berubah menjadi 301 data. hasil ekstraksi model LRFM dapat dilihat pada Tabel 1.

Tabel 1. Ekstraksi Model LRFM

CustomerID	L	R	F	M
82019001	99	114	4	1035000
82019002	0	212	1	120000
82019003	21	191	2	400000
82019004	115	96	8	1400000

4.2 *Data Preprocessing*

Hasil dari p*reprocessing data* melalui 4 tahap. Yang pertama, melakukan *selection data* yaitu pemilihan variabel dari data pelanggan yang akan digunakan proses *clustering*, lalu *transformation data* menggunakan jenis *transformation logarithmic*. Menurut penelitian (Dharmawan, 2019), dilakukan penambahan nilai 1 pada variabel LRFM yang masih memiliki nilai mendekati 0. Hal ini dilakukan untuk mengubah data agar konsisten atau memiliki jarak yang tidak terlalu jauh pada tiap variabelnya. Tabel 2 adalah hasil dari *Transformation Logarithmic LRFM* .

Tabel 2. Hasil Transformation Logartihmic LRFM

CustomerID	llog	rlog	flog	mlog
82019001	2.000	2.057	0.699	6.015
82019002	0.000	2.326	0.301	5.079
82019003	1.342	2.281	0.477	5.602
82019004	2.064	1.982	0.954	6.146

Tahap selanjutnya adalah *cleaning data* menggunakan script Klodian Dhana (2019). Hasilnya adalah terdapat 2 data yang mengandung *outlier* pada variabel *monetary* dan segera dilakukan penghapusan data tersebut. Sehingga, data yang semula 301 data menjadi 299 data. Tabel 3. Adalah hasil *cleaning data*.

Tabel 3. Hasil Cleaning Data

CustomerID	llog	rlog	flog	mlog
82019010	2.072	1.968	1.000	NA
102019011	1.833	2.009	1.000	NA
82019005	0.301	2.320	0.477	5.439
82019006	1.944	2.009	0.845	5.929
82019007	1.826	2.161	0.477	5.732
82019008	2.037	2.013	0.845	6.125
82019009	1.732	2.196	0.602	5.607

Tahap terakhir dari *preprocessing data* adalah *Normalization data.* metode yang digunakan adalah Min-Max. yaitu mengubah nilai data normalisasi menjadi skala 0 sampai 1. Dilakukan pengurangan nilai pada variabel recency. Pengurangan dengan nilai 1 pada nilai variabel recency. Dikarenakan sifat variabel *recency* berbanding terbalik dengan ketiga

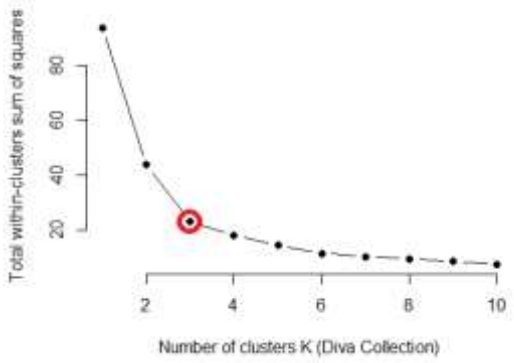
variabel lainnya (Angelie, 2017). Tabel 4. Adalah hasil normalisasi Min-Max

Tabel 4. Normalisasi Min-Max

CustomerID	norm L	R balik	norm F	norm M
82019001	0.969	0.697	0.609	0.856
82019002	0.000	0.000	0.000	0.208
82019003	0.650	0.117	0.270	0.570
82019004	1.000	0.889	1.000	0.947
82019005	0.146	0.016	0.270	0.458

4.3 Clustering Fuzzy C-Means

Sebelum memasuki tahap *clustering* metode *FCM*, terlebih dahulu menentukan acuan *cluster* optimal menggunakan *Elbow Mehod*. Hasil dari *Elbow Method* yang optimal seperti pada Gambar 2.

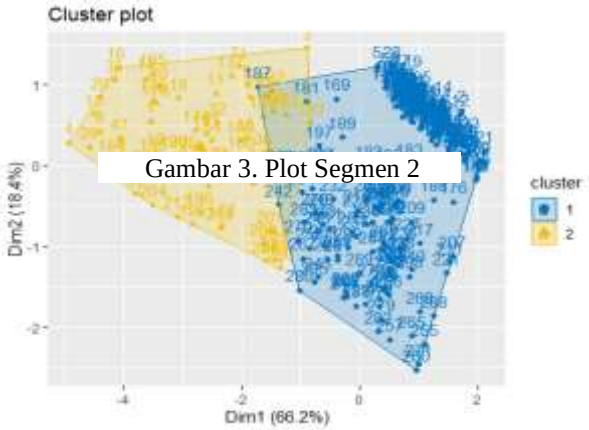


Gambar 2. Hasil Metode *Elbow*

Tahap *clustering* terbagi menjadi beberapa proses di gambar 2, yaitu penentuan jumlah *cluster* dengan *elbow method* dan implmentasi metode *clustering* dengan *fuzzy c-means*. Hasil jumlah *cluster* ditunjukkan melalui grafik pada Gambar 2. Penurunan drastis terlihat dari *K=2* ke K=3, selanjutnya penurunan secara perlahan – lahan dari K=3 menuju K=4. Dimana K=3 merupakan jumlah *cluster* optimum yang dijadikan acuan untuk proses *clustering* selanjutnya.

Selanjutnya, hasil implementasi *Fuzy C-Means* menggunakan *cluster 2* dan *cluster 3* sebagai perbndingan berdasarkan metode *elbow*. *Cluster 2* atau segmen 2 menghasilkan 63 data untuk *cluster* 1 dan 236 data untuk *cluster 2*. Gambar 3. Adalah penyebaran plot data segmen 2.

Pada gambar di atas, *cluster* 1 diagambarkan dengan plot data berwarna biru dan *cluster* 2 digambarkan dengan plot data berwarna kuning. Data yang didapat dari hasil *clustering* segmen 2 adalah nilai titik centroid pada Tabel 5, nilai derajat keanggotaan pada Tabel 6. Data dikatakan masuk ke dalam cluster apabila nilai



Gambar 3. Plot Segmen 2

semakin mendekati nilai 1. Nilai *Root Mean Squared Deviations* (RMSD) pada segmen 2 adalah 0,5802089. Dimana nilai tersebut merupakan standar deviasi dari perbedaan antara prediksi dengan nilai yang diamati. Untuk *Mean Absolute Deviations* (MAD) atau jarak rata rata tiap centroid dengan titik data adalah 3,080251. Hasil *clustering* segmen 2 (*cluster 1* dan *cluster 2)* pada Tabel 5. Hasil dari Variabel *recency* sudah dilakukan pengurangan nilai 1 dengan *recency* awal.

Tabel 5. Hasil Centroid Segmen 2

Centroid	L	R	F	M
Cluster 1	0.005	0.310	0.038	0.394
Cluster 2	0.699	0.681	0.521	0.703

Tabel 6. Derajat Keanggotaan Segmen 2

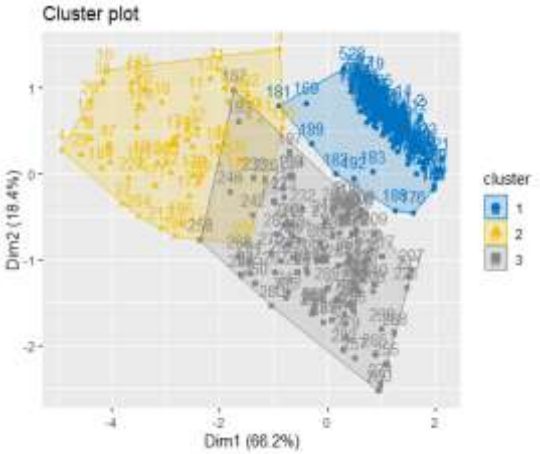
C1	C2	C3	Member_Degree
0.939	0.061	0.939	1
0.083	0.917	0.917	2
0.573	0.427	0.573	1
0.858	0.142	0.858	1
0.159	0.841	0.841	2

Tabel 7. Hasil Segmen 2

ID	L	R	F	M	Cluster
82019001	0.97	0.70	0.61	0.86	1
82019002	0.00	0.00	0.00	0.21	2
82019003	0.65	0.12	0.27	0.57	1
82019004	1.00	0.89	1.00	0.95	1
82019005	0.15	0.02	0.27	0.46	2

Hasil implementasi segmen 3 beranggotakan 3 kelompok. Yaitu *cluster 1* dengan jumlah anggota sebanyak 145 data, *cluster 2* sejumlah 55 data, dan *cluster* 3 sejumlah 99 data. Hasil segmen 3 juga digambarkan dengan penyebaran plot data seperti pada gambar 4.

Gambar 4. Adalah penyebaran data pelanggan pada segmen 3, *cluster* 1 digambarkan dengan penyebaran data berwarna biru, *cluster 2* berwarna kuning , sedangkan *cluster* 3 berwarna abu abu. Data yang didapat dari segmen 3 adalah informasi mengenai titik



centroid *cluster* pada Tabel 6, data dimasukkan ke dalam cluster apabila mendekati nilai 1 sesuai derajat keanggotaan pada Tabel 7., *Root Mean Squared Deviations* (RMSD) atau standar deviiasi dari perbedaan nilai kenyataan dengan nilai prediksi sebelumnya, sebesar 0,7488091, *Mean Absolute Deviation* (MAD) sebesar 4,373383, dan untuk hasil akhir *clustering* segmen 3 terdapat pada Tabel 8.

Tabel 8. Hasil Centroid Segmen 3

Centroid	L	R	F	M
Cluster 1	0.004	0.154	0.011	0.390
Cluster 2	0.830	0.681	0.583	0.744
Cluster 3	0.009	0.715	0.126	0.440

Tabel 9. Derajat Keanggotaan Segmen 3

C1	C2	C3	Member_Degree
0.017	0.959	0.024	2
0.885	0.028	0.086	1
0.367	0.397	0.236	2
0.082	0.804	0.114	2
0.763	0.077	0.160	1

Tabel 10. Hasil Segmen 3

ID	L	R	F	M	Cluster
82019001	0.97	0.70	0.61	0.86	2
82019002	0.00	0.00	0.00	0.21	1
82019003	0.65	0.12	0.27	0.57	2
82019004	1.00	0.89	1.00	0.95	2
82019005	0.15	0.02	0.27	0.46	1

4.4 Uji Validitas Cluster

Uji Validitas Cluster menggunakan 3 jenis pengujian. Yaitu Uji *Partition Coefficient (PC), clValid (Connectivity, Dunn Index,* dan *Silhoutte), dan Davies Bouldin's Index*. Hasil uji pertama, *Partition Coefficient (PC)* pada segmen 2 senilai 0,804234482181102 dan segmen 3 senilai 0,74633087868498. Analisa yang didapatkan, segmen 2 lebih baik daripada segmen 3. Hal ini dikarenakan semakin tinggi nilai PC, maka data cenderung tidak monoton terhadap jumlah cluster.

SEGMEN 2
(CLUSTER 1 dan CLUSTER2

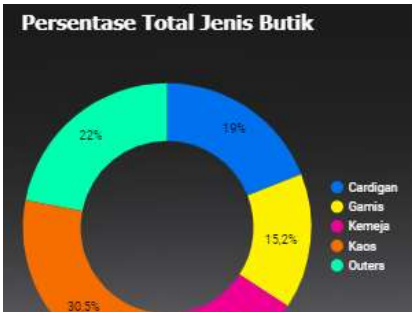
Implementasi uji yang kedua, menggunakan uji *clValid*. Terdapat tiga jenis pengujian di dalamnya, yaitu *Connectivity, Dunn Index,* dan *Silhoutte*. Pengujian diuji dari segmen 2 hingga segmen 10. Hasil *Connectivity* pada segmen 2 senilai 6.267460317 dan segmen 3 senilai 16.1484127. pada Dunn Index, segmen 2 memiliki nilai 0.171267489 dan segmen 3 senilai 0.01983532. sedangkan pada Silhoutte, segmen 2 memiliki nilai 0.571747658 dan segmen 3 dengan nilai 0.533836269. hasil analisa yang didapatkan pada uji *Connectivity*, segmen 2 lebih baik daripada segmen 3. Hal ini dikarenakan apabilai nilai *connectivity* mendekati angka 0, maka *cluster* semakin tepat. Pada *Dunn Index*, apabila semakin tinggi nilai *Dunn Index*, maka rasio suatu jarak antar data akan semakin mengecil di satu kelompok yang sama terhadap kelompok yang berbeda . Hasil dari *Silhoutte*, semakin tinggi nilai *Silhoutte*, maka *cluster* menunjukkan posisi yang optimal. Hasil analisa yang dapat disimpulkan pada uji clValid, segmen 2 merupakan segmen yang

terbaik jika dibandingkan dengan segmen lainnya yang telah diuji.

Pengujian ketiga menggunakan uji Davies Bouldin's Index. Pengujian ini dilakukan pada segmen 2 hingga segmen 10. Hasil dari segmen 2 adalah 0.8152322 dan segmen 3 adalah 0.8217698. Segmen 2 merupakan segmen paling baik daripada segmen lainnya yang telah diuji. Hal ini dikarenakan jika nilai DBI semakin rendah, maka jarak suatu data dengan data lainnya akan semakin tepat. Nilai segmen 2 berada pada nilai paling rendah.

4.5 Visualisasi Dashboard

Hasil dari analisa *clustering Fuzzy C-Means* divisualisasikan dalam bentuk *dashboard*. Data digambarkan dalam bentuk diagram batang, lingkaran, serta tabel informasi. Visualisasi *dashboard* dibuat menggunakan *Google Data Studio*. Halaman pertama *dashboard* berisikan informasi mengenai Toko Butik Diva Collection secara keseluruhan, halaman kedua berisi data *clustering* segmen 2, dan halaman ketiga berisi data *clustering* segmen 3.



Gambar 4. Persentase Total Jenis Butik

Pada halaman pertama, terdapat informasi profil butik, menu drop-down untuk memilih Periode dan Tanggal, diagram kartu skor yang menampilkan jumlah transaksi dan total pendapatan, diagram batang, persentase butik diagram lingkaran dan tabel informasi pelanggan bersasarkan periode transaksi yang telah dipilih sebelumnya, serta informasi pelanggan. secara keseluruhan berdasarkan periode bulan Agutsus 2019 – Desember 2019, Butik Diva Collection berhasil menjual 1330 unit baju dan bulan September adalah penjualan terbanyak dengan jumlah 490 unit baju yang terjual. Rata rata pelanggan melakukan 3 kali transaksi selama periode tersebut dan baju ukuran S adalah penjualan terbanyak dengan jumlah 475 baju.

Pada halaman ke-2, menampilkan menu drop-down untuk memilih periode waktu, *slider*



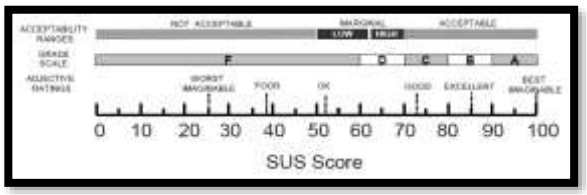
Gambar 3. Menu Input Segmen 2

kontrol untuk memilih jumlah nilai masing masing variabel LRFM, tabel informasi dan diagram batang dari data yang telah diinputkan sebelumnya. Halaman ke-2 terdapat tabel informasi total unit baju yang terjual dan total pendapatan berdasarkan cluster 1 dan cluster 2 yang berada di segmen 2. Total penjualan jenis baju periode dan total pendapatan uang Agustus 2019 hingga Desember 2019 berdasarkan cluster di segmen 2, digambarkan dengan diagram batang.

Halaman ke-3 menampilkan hasil dari *clustering* segmen 3 yang berisi *cluster* 1, *cluster* 2, dan *cluster* 3. Terdapat menu drop-down yang bisa dipilih oleh *user*, begitu juga dengan nilai masing masing variabael LRFM menggunakan *slider* kontrol. Nilai yang sudah dimasukkan oleh *user* akan menghasilkan informasi berupa tabel dan diagram batang. Tabel berupa informasi cluster transaksi pelanggan, perbandingan total frekuensi dan perbandingan total jenis baju yang terjual di segmen 3. Terdapat tabel yang menginformasikan peringkat *cluster* terbaik hingga terburuk di segmen 3. Pemeringkatan berdasarkan total unit baju yang terjual dan total pendapatan.

4.6 *System Usability Scale*

Pengujian kualitas dari *dashboard* yang dibuat, diuji menggunakan *Usability Testing.* Jenis *System Usability Testing* yang akan digunakan untuk melakukan uji *dashboard.* Penguji *dashboard* adalah 3 responden dari pihak Butik Diva Collection. Yaitu pemilik, manajer, dan pemegang keuangan Butik Diva Collection. Total nilai dari responden pertama adalah 77.5, responden ke-dua adalah 75, dan responden ke-3 adalah 80. Nilai rata-rata dari ketiga responden adalah 77.5.



Gambar 5. Penentuan Nilai Hasil

Menurut (Bangor *et al.*, 2009), berdasarkan hasil SUS yang dioleh dari kuesioner Butik Diva Collection, nilai rata-rata yang didapatkan yaitu 77.5, nilai tersebut termasuk ke dalam kategori *acceptable*. Hasil *dashbaord* tersebut dapat diterima dengan baik serta dapat digunakan dengan mudah dalam memahami perilaku pelanggan Toko Butik Diva Collection. Untuk peringkat nilai denan skor 77.5, termasuk ke dalam peringkat 'C'. dimana jika dilihat dari *adjective rating*, nilai tersebut masuk ke dalam kategori 'GOOD'.

5. Kesimpulan

Perilaku atau sifat setiap pelanggan dapat dilihat dari hasil *clustering Fuzzy C-Means* berdasarkan tiap-tiap variabel LRFM. proses yang dilakukan adalah ekstraksi model LRFM, data *preprocessing*, *clustering* menggunakan metode *Fuzzy C-Means*, Uji validitas dan verifikasi *cluster*, visualisasi dashboard dan pengujian *Usability Testing.* Pada *preprocessing data*, terdapat 560 data mentah pelanggan yang siap diolah di tahap *cleaning data*, *recode variable*, *calculate* LRFM, *transformation*, *cleaning outlier*, dan normalisasi Min-Max dan menjadi 299 data untuk proses *clustering Fuzzy C-Means*. Sebelum melakukan *clustering*, dilakukan proses pencarian jumlah *cluster* yang optimal menggunakan Metode Elbow untuk mencari acuan jumlah *cluster* yang optimum, yaitu ditemukan *cluster* K=3. Jumlah 3 cluster yang akan dijadikan acuan. Untuk implementasi *clustering*, menggunakan segmen 2 dan segmen 3 sebagai perbandingan *cluster.*

Hasil uji validitas dan verifikasi menunjukkan segmen *2* adalah segmen paling optimal. Hal ini berdasarkan kesesuaian derajat keanggotaan data dengan jarak data terhadap titik *centroid* cluster. Di dalam segmen 2 terdapat 2 *cluster*. Yaitu *cluster* 1 dan *cluster* 2. Uji validitas menggunakan metode Partition Coefficient (PC), Davies Bouldin's Index, dan clValid (Connectivity, Dunn Index, dan Silhoutte). Hasil dari seluruh pengujian validitas tersebut menghasilkan *cluster* 2 yang paling optimum. Hasil ini berbeda dengan hasil Metode Elbow yang menunjukkan jumlah 3 *cluster* yang optimum. Oleh karena itu, dilakukan 2 kali proses *clustering* dengan jumlah 2 *cluster*.

Penggambaran hasil perilaku pelanggan yang sudah didenormalisasi dari proses *clustering* diimplementasikan dalam bentuk *dashboard* yang informatif. Dimana pihak Butik Diva Collection dapat melihat perilaku pelanggan dalam bentuk diagram batang, diagram lingkaran, tabel informasi, pemilihan menu drop-down, dan *slider* kontrol untuk mempermudah analisa hasil *clustering*. Terdapat 3 halaman. Halaman pertama menjelaskan profil butik diva colletion,halaman kedua menggambarkan segmen 2 dan halaman ketiga menggambarkan segmen 3. Pihak Butik Diva Collection melakukan pengujian visualisasi *dashboard* menggunakan *Usability Testing* jenis *System Usability Scale*. Hasilnya menunjukkan nilai 77.5 dari 3 responden. Nilai tersebut termasuk ke dalam kategori *acceptable* dan *grade scale* bernilai 'C'. nilai tersebut menunjukkan bahwa *dashboard* tersebut dapat diterima dengan baik dan digunakan dengan mudah dalam menganalisa perilaku pelanggan oleh pihak Butik Diva Collection.

6. DAFTAR PUSTAKA

Angelie, A. V. (2017) *Segmentasi Pelanggan Menggunakan Clustering K-Means Dan Model Rfm (Studi Kasus : PT . Bina Adidaya Surabayaa) Customer Segmentation Using K-Means Clustering and Rfm Model (Case Study : Pt . Bina Adidaya Surabaya Segmentasi Pelanggan Menggunakan Clusteri*.

Bangor, A. *et al.* (2009) 'Determining what individual SUS scores mean: adding an adjective rating scale', *Journal of usability studies*, 4(3), pp. 114–123.

Brooke, J. (2020) 'SUS: A "Quick and Dirty" Usability Scale', *Usability Evaluation In Industry*, (January 1996), pp. 207–212. doi: 10.1201/9781498710411-35.

Buttle, F. (2013) *Customer Relationship Management Concepts And Technologies*, *Journal of Chemical Information and Modeling*. doi: 10.1017/CBO9781107415324.004.

Dharmawan, M. T. (2019) 'Segmentasi Pelanggan Menggunakan Metode Fuzzy

C-Means Clustering Berdasarkan LRFM Model Pada Toko Sepatu (Studi Kasus : Ride Inc Kota Malang)', *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3(2), pp. 1978–1985.

Efrain Turban, Jay E. Aronson, T. P.-L. (2005) *Decision Support Systems and Intelligent Systems Jilid I Edisi 7*. 7th edn. Edited by Andi. Yogyakarta: Andi. doi: 09101.196.

Han, K. dan (2006) 'Data Mining: Data Mining Concepts and Techniques', in *Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013*. 3rd edn, pp. 448–450. doi: 10.1109/ICMIRA.2013.45.

Larose, D. T. (2005) *DISCOVERING KNOWLEDGE IN DATA An Introduction to Data Mining Second Edition Wiley Series on Methods and Applications in Data Mining*.

Mahdulatha, T. S. (2012) 'An Overview of Clustering Methods', *IOSR Journal of Engineering*, 11(6), pp. 719–725. doi: 10.9790/3021-0204719725.

Rahakbauw, D. L., Ilwaru, V. Y. I. and Hahury, M. H. (2017) 'Implementasi Fuzzy C-Means Clustering Dalam Implementation Of Fuzzy C-Means Clustering In', 11, pp. 1–12.

Wu, H. H., Lin, S. Y. and Liu, C. W. (2014) 'Analyzing patients' values by applying cluster analysis and LRFM model in a pediatric dental clinic in Taiwan', *Scientific World Journal*, 2014. doi: 10.1155/2014/685495.