

SISTEM PREDIKSI *CUSTOMER LOYALTY* DENGAN METODE RFM DAN FUZZY C-MEANS

David¹, Manatap Dolok Lauro², Dyah Erny Herwindiati³

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara,
Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

E-mail:¹davidpheng2905@gmail.com, ²manataps@fti.untar.ac.id, ³dyahh@fti.untar.ac.id

ABSTRAK

Tulisan ini membahas tentang customer loyalty dengan menggunakan metode RFM dan Fuzzy C-Means. RFM adalah kepanjangan dari Recency, Frequency dan Monetary. Yang dimaksud dengan Recency adalah rentang tanggal transaksi yang dilakukan pelanggan, sedangkan Frequency adalah jumlah transaksi yang dilakukan pelanggan selama periode tertentu, dan Monetary adalah jumlah nominal uang yang dikeluarkan selama periode tertentu. Data diambil dari Toko Cianjur dengan riwayat transaksi dari Januari 2018 hingga Desember 2018 yang terdiri dari 892 transaksi. Fuzzy C-Means adalah salah satu algoritma clustering yang menerapkan pendekatan fuzzy untuk menentukan cluster berdasarkan degree of membership. Partition Coefficient Index dan Partition Entropy Index digunakan sebagai metode evaluasi. Tujuan pembuatan sistem prediksi customer loyalty ini yaitu mengetahui pelanggan manakah yang memiliki loyalitas yang tinggi terhadap toko sehingga dapat menjadi bahan pertimbangan toko untuk memberikan diskon kepada pelanggan. Hasil pengujian yang didapat menyatakan bahwa hasil cluster pembentukan dua klaster dengan $m = 2$ menunjukkan bahwa pelanggan cukup sering berbelanja terhadap toko sampai Desember 2018. Sedangkan hasil pembentukan tiga klaster dengan $m = 2$ menunjukkan bahwa pelanggan sering melakukan transaksi serta nilai monetary yang cukup besar terhadap toko.

Kata kunci—Clustering, Customer Loyalty, Fuzzy C-Means, RFM

ABSTRACT

This paper discusses about customer loyalty using RFM method and Fuzzy C-Means. RFM stands for Recency, Frequency, and Monetary. What is meant by Recency is the date range of transactions made by customers, while Frequency is the number of transactions made by customers during a certain period, and Monetary is the nominal amount of money spent during a certain period. Data is taken from Toko Cianjur with transactions history from January 2018 to December 2018 consisting of 892 transactions. Fuzzy C-Means is one of clustering algorithm which applies fuzzy approach to determine the clusters based on the degree of membership. Partition Coefficient Index and Partition Entropy Index are used as the evaluation methods. The goals of making this application that the store can find customers who have a high loyalty value to store and can be used as consideration for store to provide discounts to customers. The test results obtained state that the results of the cluster formation of two clusters with $m = 2$ indicate that customers quite often shop at stores until December 2018. While the results of the formation of three clusters with $m = 2$ indicate that customers often make transactions and a large monetary value to the store.

Keywords—Clustering, Customer Loyalty, Fuzzy C-Means, RFM

1. PENDAHULUAN

Semakin maju dan berkembangnya dunia usaha dengan pesat, menciptakan tingkat persaingan yang semakin ketat antar toko. Toko Cianjur merupakan toko yang bergerak dalam bidang penjualan barang elektronik dan furnitur. Masalah yang dihadapi oleh toko saat ini adalah sulitnya mengetahui pelanggan yang memiliki nilai loyalitas yang tinggi terhadap toko. Penyelesaian terhadap masalah yang dihadapi toko adalah dengan membuat Aplikasi Prediksi *Customer Loyalty*. Aplikasi Prediksi *Customer Loyalty* merupakan aplikasi yang dibuat dengan menggunakan Metode RFM dan *Fuzzy C-Means*. Studi kasus yang digunakan dalam pembuatan aplikasi ini adalah data transaksi yang dilakukan oleh pelanggan terhadap toko selama periode Januari 2018 hingga Desember 2018 dengan data pelanggan sebanyak 578 data dan 892 data transaksi.

Metode yang digunakan untuk mengukur *customer loyalty* adalah metode RFM. Metode RFM mampu mengatasi pengelolaan data yang kurang baik karena data transaksi pelanggan cenderung memiliki banyak atribut. Metode RFM merupakan kepanjangan dari *Recency*, *Frequency*, dan *Monetary*. Yang dimaksud dengan *Recency* adalah rentang tanggal transaksi yang dilakukan pelanggan, sedangkan *Frequency* adalah jumlah transaksi yang dilakukan pelanggan selama periode tertentu, dan *Monetary* adalah jumlah nominal uang yang dikeluarkan selama periode tertentu. Metode RFM menggunakan data riwayat transaksi yang dilakukan oleh pelanggan.

Pada tulisan ini akan dikelompokkan juga kesetiaan pelanggan menjadi 3 kelompok yaitu tidak loyal, loyal, dan sangat loyal. Metode pengelompokan yg digunakan adalah Fuzzy C-Means. *Clustering Fuzzy C-Means* digunakan untuk mengelompokkan data dimana keberadaan tiap titik data dalam suatu *cluster* ditentukan oleh *membership degree*. Fuzzy C-Means mampu menempatkan suatu data yang terletak diantara dua atau lebih *cluster* yang lain pada suatu *cluster*, sehingga data dapat menjadi anggota dari semua kelas yang terbentuk dengan *membership degree* yang berada diantara 0 hingga 1. Metode Fuzzy C-Means menggunakan derajat keanggotaan tiap-tiap titik data untuk menghitung pusat kelompok data secara berulang-ulang untuk masing-masing *cluster* hingga pusat kelompok akan bergerak menuju lokasi yang tepat. Evaluasi yang digunakan untuk menguji hasil klaster yang terbentuk menggunakan evaluasi *Coefficient of Variation*, *Partition Coefficient Index* dan *Partition Entropy Index*.

Tujuan dari melakukan penelitian ini adalah untuk mengetahui pelanggan manakah yang memiliki tingkat loyalitas yang tinggi terhadap toko sehingga dapat menjadi bahan pertimbangan untuk toko dalam menentukan strategi yang tepat terhadap kelompok yang didapat.

2. METODE PENELITIAN

2.1 Metode RFM

Model segmentasi RFM (*Recency*, *Frequency*, dan *Monetary*) telah banyak digunakan dalam berbagai bidang, terutama dalam dunia pemasaran. Model RFM dapat melacak transaksi pelanggan terhadap waktu, frekuensi dan jumlah yang dikeluarkan saat transaksi. Dengan model RFM, seorang pengambil keputusan dapat mengidentifikasi pelanggan yang berharga dan mengembangkan strategi pemasaran yang efektif. Berikut definisi dari variabel RFM[1] :

1. Variabel R (*Recency*)

Recency merupakan variabel yang mempresentasikan adanya rentang transaksi antar waktu pelanggan melakukan transaksi terakhir dengan waktu analisis yang ditentukan.

2. Variabel F (*Frequency*)

Frequency merupakan variabel yang mempresentasikan berapa kali atau jumlah pembelian yang dilakukan pelanggan selama waktu analisis yang ditentukan.

3. Variabel M (*Monetary*)

Monetary merupakan variabel yang mempresentasikan total biaya transaksi yang dilakukan pelanggan selama waktu analisis terhadap toko. Nilai *monetary* didapat dari perhitungan proses pembayaran yang dilakukan pelanggan.

2.2. Clustering

Clustering atau biasa disebut dengan klasterisasi adalah metode pengelompokan data tanpa berdasarkan kelas data yang telah ditentukan[2]. *Clustering* termasuk metode *unsupervised learning*, yang artinya klasterisasi ini tidak memiliki data latih. Memaksimalkan kemiripan antar anggota dalam suatu kelas dan meminimumkan kemiripan antar klaster merupakan tujuan dari klasterisasi. Pemilihan pusat klaster yang kurang baik dengan cara acak merupakan kelemahan dari metode clustering. Algoritma klasterisasi membutuhkan jarak untuk mengukur kemiripan antar data.

2.3. Logika Fuzzy

Lotfi Zadeh memperkenalkan logika *fuzzy* dan teori himpunan *fuzzy* sebagai cara mengatasi bentuk masalah ketidakpastian. Kata *fuzzy* merupakan kata sifat yang memiliki arti kabur atau samar. Logika *fuzzy* didasarkan pada ketidakpastian batas antara suatu kriteria dengan kriteria lainnya yang disebabkan adanya penilaian manusia terhadap sesuatu. Logika *fuzzy* menggunakan ungkapan bahasa untuk menggambarkan nilai variabel. Logika *fuzzy* bekerja dengan menggunakan *membership degree* dari sebuah nilai yang kemudian digunakan untuk menentukan hasil yang ingin dihasilkan berdasarkan atas spesifikasi yang ditentukan. Nilai keanggotaannya menunjukkan bahwa suatu variabel tidak tertuju pada 0 (salah) atau 1 (benar), melainkan nilai yang mempunyai selang di antara 0 sampai dengan 1[3].

2.4. Metode Fuzzy C-Means

Fuzzy C-Means merupakan salah satu teknik untuk menentukan *cluster* data dimana posisi tiap-tiap titik data dalam suatu *cluster* ditentukan oleh *membership degree*. FCM menggunakan model pengelompokan *fuzzy* dengan indeks kekaburan sehingga data dapat menjadi anggota dari semua kelas atau *cluster* yang terbentuk dengan *membership degree* yang berada di antara 0 hingga 1[4].

Konsep dasar FCM, pertama kali adalah menentukan pusat klaster, yang akan menandai lokasi rata-rata untuk setiap klaster. Pada awal kondisi, pusat klaster masih belum akurat. Tiap-tiap titik data memiliki *membership degree* untuk tiap klaster yang terbentuk. Dengan cara memperbaiki pusat kelompok dan nilai keanggotaan tiap-tiap data secara berulang, maka akan dapat dilihat bahwa pusat kelompok akan bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi objektif yang menggambarkan jarak dari titik data yang diberikan ke pusat klaster yang terbobot oleh *membership degree* data tersebut. Keunggulan Fuzzy C-Means adalah dapat memberikan hasil pengelompokan bagi objek-objek yang tersebar tidak teratur. Output dari FCM merupakan derajat pusat *cluster* dan beberapa *membership degree* untuk tiap titik data.

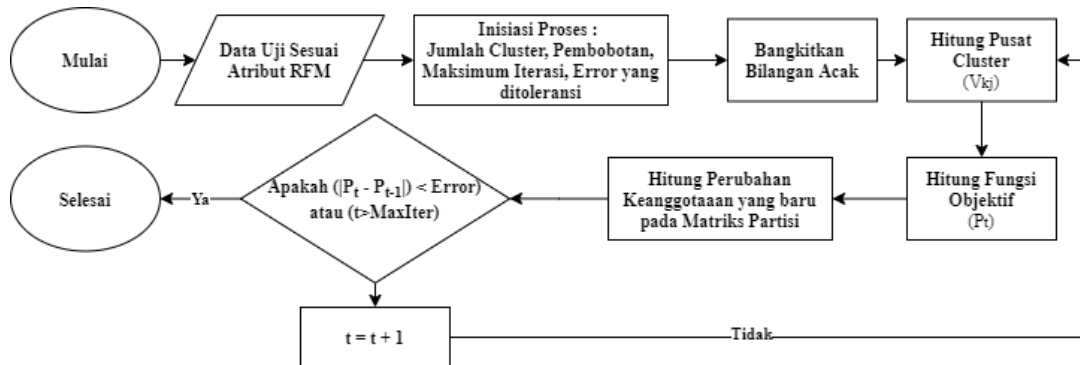
Berikut algoritma *Fuzzy C-Means* :

1. Masukkan data yang akan di *cluster* kedalam sebuah matriks X, dimana matriks berukuran $m \times n$, dengan m adalah jumlah data yang akan diklaster dan n adalah atribut setiap data. Contoh X_{ij} = data ke- i ($i=1,2,...,m$), atribut ke- j ($j=1,2,...,n$).

2. Batasan :

- a. Jumlah *cluster* = c ;
- b. Pembobotan atau *fuzziness*($w > 1$) = w ;
- c. Maksimum iterasi = MaxIter ;
- d. Error yang ditoleransi = ε ;
- e. Fungsi Objektif Awal(P_0) = 0;
- f. Iterasi Awal(t) = 1;

Jumlah cluster, pembobotan, maksimum iterasi dan error yang ditoleransi ditentukan oleh pengguna. Pada algoritma FCM, nilai w tidak memiliki nilai ketepatan, biasanya nilai $w > 1$ dan pada umumnya diberi nilai 2 [5]. Semakin besar nilai w maka akan semakin besar pula pembobotan dan klaster semakin kabur.



Gambar 1 Flowchart Algoritma Fuzzy C-Means

3. Bangkitkan bilangan acak μ_{ik} (dengan $i=1,2,\dots,m$ dan $k=1,2,\dots,c$) sebagai elemen matriks partisi awal U , dengan X_i adalah data ke- i

$$U = \begin{bmatrix} \mu_{11}(X_1) & \mu_{21}(X_1) & \dots & \mu_{c1}(X_1) \\ \mu_{12}(X_2) & \mu_{21}(X_2) & \dots & \mu_{c2}(X_2) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{1i}(X_i) & \mu_{2i}(X_i) & \dots & \mu_{ci}(X_i) \end{bmatrix} \quad (1)$$

Dengan jumlah kolom dalam satu baris adalah 1(satu). μ_{ik} adalah *membership degree* yang merujuk pada seberapa besar kemungkinan suatu data dapat menjadi anggota ke dalam suatu *cluster*.

$$\sum_{i=1}^c \mu_{ci} = 1 \quad (2)$$

4. Hitung pusat *cluster* ke- k : V_{kj} , dengan $k=1,2,\dots,c$ dan $j=1,2,\dots,n$

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n ((\mu_{ik})^w)} \quad (3)$$

5. Hitung fungsi objektif pada iterasi ke- t , P_t :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (4)$$

Fungsi objektif digunakan sebagai syarat perulangan untuk mendapatkan pusat *cluster* yang tepat. Sehingga diperoleh kecenderungan data akan masuk ke *cluster* mana pada langkah akhir.

6. Hitung perubahan *membership degrees* setiap data pada *cluster* (perbaiki matriks partisi U) dengan :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}; i = 1, 2, 3, \dots, n; k = 1, 2, 3, \dots, c \quad (5)$$

7. Cek kondisi berhenti :
- Jika $(|P_t - P_{t-1}| < \varepsilon)$ atau $t > MaxIter$ maka berhenti;
 - Jika tidak : $t=t+1$, ulangi langkah 4.

2.5. Coefficient of Variation

Coefficient of Variation (CV) merupakan sistem pada sebuah perbandingan yakni antara simpangan yang standar serta nilai hitung rata - rata yang dapat dinyatakan dalam bentuk sebuah persentase[6]. Koefisien variasi didefinisikan sebagai rasio dari standar deviasi σ dengan mean μ :[7]

$$C_v = \frac{\sigma}{\mu} \quad (6)$$

dan estimasi dari koefisien variasi digunakan :

$$\hat{C}_v = \frac{s}{\bar{x}} \quad (7)$$

$$\text{dengan } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \text{ dan } s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (8)$$

Nilai dari koefisien variasi yang lebih kecil memiliki arti bahwa data yang didapat lebih seragam (homogen), sebaliknya jika nilai dari koefisien variasi lebih besar memiliki arti bahwa data yang didapat lebih bervariasi (heterogen).

2.6. Partition Coefficient Index

Bedzek(1981) mengusulkan validasi fuzzy clustering dengan menghitung partisi atau PC sebagai evaluasi nilai keanggotaan data pada tiap *cluster*. Nilai PCI hanya melakukan evaluasi terhadap nilai *membership degree*. Nilai PCI memiliki rentang 0 hingga 1, nilai yang mendekati angka 1(satu artinya mempunyai kualitas kluster yang didapat semakin baik. Berikut rumus validasi PCI :[8]

$$PCI = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k \mu_{ij}^2 \quad (9)$$

Keterangan :

N = jumlah data dalam set data

K = jumlah kluster

μ_{ij} = nilai keanggotaan data ke-i pada kluster ke-j

2.7 Partition Entropy Index

Indeks validitas yang pertama kali berhubungan dengan algoritma Fuzzy C-Means adalah Partition Entropy Index (Bezdek, 1974) didefinisikan dengan persamaan berikut :[9]

$$PEI = -\frac{1}{n} \sum_{i=1}^c \sum_{j=1}^n \mu_{ij} \log(\mu_{ij}) \quad (10)$$

Partition Entropy merupakan evaluasi yang mengukur tingkat kekaburan dari partisi cluster. Pada umumnya, kluster optimal didapatkan jika nilai yang diperoleh mendekati kecil (mendekati 0).

3. HASIL DAN PEMBAHASAN

Hasil pengujian dari sistem prediksi *customer loyalty* dengan metode RFM dan *Fuzzy C-Means* ini terdiri dari 2 tahap pengujian, yaitu:

1. Pengujian Terhadap Modul

Pengujian terhadap modul bertujuan untuk mengecek apakah modul – modul dalam aplikasi sudah dapat berjalan dengan baik. Hasil pengujian menunjukkan bahwa semua modul telah berjalan dengan baik dan sesuai dengan rancangan. Berikut modul – modul yang terdapat dalam aplikasi :

- a. Modul Login
- b. Modul Home
- c. Modul Pelanggan
- d. Modul Transaksi
- e. Modul Klaster
- f. Modul Riwayat Klasterisasi
- g. Modul Pengaturan
- h. Modul Tentang

2. Pengujian Terhadap Kinerja Program

Pengujian terhadap kinerja program bertujuan untuk mengetahui apakah program sudah bekerja dengan baik sesuai dengan dasar teori yang telah dijabarkan diatas. Proses pengujian dilakukan dengan cara membandingkan hasil perhitungan nilai evaluasi *Partition Coefficient Index* dan *Partition Entropy Index* pada program dengan hasil perhitungan secara manual. Hasil yang sama menunjukkan bahwa program sudah berjalan dengan baik sesuai dengan teori. Hasil dapat dilihat pada tabel 1.

Tabel 1 Evaluasi *Partition Coefficient Index* dan *Partition Entropy Index*

Cluster	PCI (Program)	PCI (Manual)
3	0.709805	0.709805
Cluster	PEI (Program)	PEI (Manual)
3	0.233599	0.233599

Pengujian selanjutnya dilakukan dengan evaluasi terhadap klaster yang terbentuk. Evaluasi yang digunakan adalah *Coefficient of Variation*, *Partition Coefficient Index*, dan *Partition Entropy Index*.

3.1 *Coefficient of Variation*

Pengujian dilakukan terhadap data transaksi yang dilakukan oleh pelanggan terhadap toko selama periode Januari 2018 hingga Desember 2018. Pengujian dilakukan terhadap pembentukan dua klaster, tiga klaster, dan empat klaster dengan inputan pembobotan yang berbeda yakni 2 dan 3.

1. Dua Klaster (Pembobotan = 2)

Hasil evaluasi terhadap dua klaster didapat bahwa nilai *coefficient of variation* dari *cluster 1* (C1) adalah data lebih heterogen dengan nilai rata-rata RFM yaitu *recency* sebesar 247, frekuensi hanya 1, *monetary* sebesar 2.024.213 serta nilai simpangan baku yang didapat adalah 1266321,88 , sedangkan *cluster 2* (C2) adalah data lebih homogen dari *cluster 1* (C1) dengan nilai rata-rata RFM yaitu *recency* sebesar 190, frekuensi sebesar 3, *monetary* sebesar 8.652.612 serta nilai simpangan baku yang didapat adalah 4118705,87.

2. Tiga Klaster (Pembobotan = 2)

Hasil evaluasi terhadap tiga klaster didapat bahwa nilai *coefficientofvariation* dari *cluster* 3 (C3) adalah data lebih heterogen dengan nilai rata-rata RFM yaitu *recency* sebesar 253, frekuensi hanya 1, *monetary* sebesar 1.486.299 serta nilai simpangan baku yang didapat adalah 773763,02, sedangkan *cluster* 1 (C1) adalah data lebih homogen dari *cluster* lain dengan nilai rata-rata RFM yaitu *recency* sebesar 215, frekuensi sebesar 2, *monetary* sebesar 4.753.052 serta nilai simpangan baku yang didapat adalah 1235755,89.

3. Empat Klaster (Pembobotan = 2)

Hasil evaluasi terhadap empat klaster didapat bahwa nilai *coefficientofvariation* dari *cluster* 3 (C3) adalah data lebih heterogen dengan nilai rata-rata RFM yaitu *recency* sebesar 254, frekuensi hanya 1, *monetary* sebesar 1.443.275 serta nilai simpangan baku yang didapat adalah 742572,24, sedangkan *cluster* 4 (C4) adalah data lebih homogen dari *cluster* lain dengan nilai rata-rata RFM yaitu *recency* sebesar 205, frekuensi sebesar 5, *monetary* sebesar 22.262.000 serta nilai simpangan baku yang didapat adalah 4395863,51.

4. Dua Klaster (Pembobotan = 3)

Hasil evaluasi terhadap dua klaster didapat bahwa nilai *coefficientofvariation* dari *cluster* 1 (C1) adalah data lebih heterogen dengan nilai rata-rata RFM yaitu *recency* sebesar 247, frekuensi hanya 1, *monetary* sebesar 1.951.792 serta nilai simpangan baku yang didapat adalah 1195731,39, sedangkan *cluster* 2 (C2) adalah data lebih homogen dari *cluster* 1 (C1) dengan nilai rata-rata RFM yaitu *recency* sebesar 193, frekuensi sebesar 3, *monetary* sebesar 8.202.650 serta nilai simpangan baku yang didapat adalah 4051523,50.

5. Tiga Klaster (Pembobotan = 3)

Hasil evaluasi terhadap tiga klaster didapat bahwa nilai *coefficientofvariation* dari *cluster* 3 (C3) adalah data lebih heterogen dengan nilai rata-rata RFM yaitu *recency* sebesar 254, frekuensi hanya 1, *monetary* sebesar 1.443.275 serta nilai simpangan baku yang didapat adalah 742572,24, sedangkan *cluster* 1 (C1) adalah data lebih homogen dari *cluster* lain dengan nilai rata-rata RFM yaitu *recency* sebesar 219, frekuensi sebesar 2, *monetary* sebesar 4.483.248 serta nilai simpangan baku yang didapat adalah 1117230,12.

6. Empat Klaster (Pembobotan = 3)

Hasil evaluasi terhadap empat klaster didapat bahwa nilai *coefficientofvariation* dari *cluster* 1 (C1) adalah data lebih heterogen dengan nilai rata-rata RFM yaitu *recency* sebesar 258, frekuensi hanya 1, *monetary* sebesar 817.447 serta nilai simpangan baku yang didapat adalah 420823,35, sedangkan *cluster* 4 (C4) adalah data lebih homogen dari *cluster* lain dengan nilai rata-rata RFM yaitu *recency* sebesar 205, frekuensi sebesar 2, *monetary* sebesar 5.194.348 serta nilai simpangan baku yang didapat adalah 1123156,94. Berikut tabel hasil pengujian evaluasi *Coefficient of Variation* (CV) dari pengujian diatas:

Tabel 2 Nilai *Coefficient of Variation*

Nilai <i>Coefficient of Variation</i>				
Pembobotan = 2	C1	C2	C3	C4
Dua Klaster	62.55%	47.60%	-	-
Tiga Klaster	25.99%	39.33%	52.05%	-
Empat Klaster	22.42%	20.41%	51.45%	19.74%
Pembobotan = 3	C1	C2	C3	C4
Dua Klaster	61.26%	49.39%	-	-
Tiga Klaster	24.92%	40.75%	51.45%	-
Empat Klaster	51.48%	39.33%	25.84%	21.62%

3.2. Partition Coefficient Index

Hasil pengujian PCI dilakukan terhadap pembentukan dua klaster, tiga klaster, dan empat klaster dengan inputan pembobotan 2 dan pembobotan 3 menunjukkan bahwa hasil inputan

pembobotan 2 menghasilkan nilai yang lebih baik dari pembobotan 3. Hasil dapat dilihat pada Tabel 3.

Tabel 3 Nilai *Partition Coefficient Index*

Nilai Partition Coefficient Index			
Pembobotan	Dua Klaster	Tiga Klaster	Empat Klaster
2	0.893393	0.841814	0.828772
3	0.759166	0.629019	0.543501

3.3 Partition Entropy Index

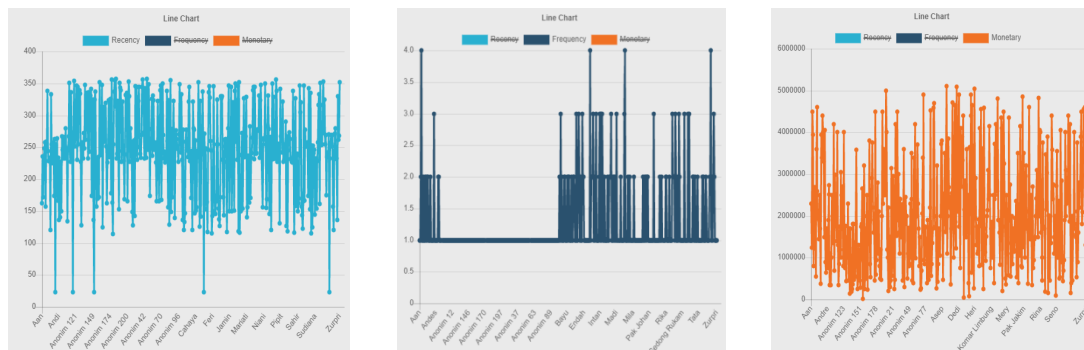
Hasil pengujian PEI dilakukan terhadap pembentukan dua klaster, tiga klaster, dan empat klaster dengan inputan pembobotan 2 dan pembobotan 3 menunjukkan bahwa hasil inputan pembobotan 2 menghasilkan nilai yang lebih baik dari pembobotan 3. Hasil dapat dilihat pada Tabel 4.

Tabel 4 Nilai *Partition Entropy Index*

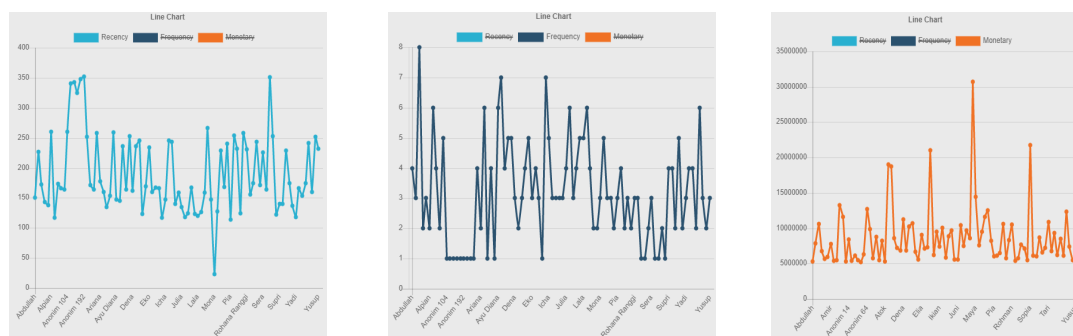
Nilai Partition Entropy Index			
Pembobotan	Dua Klaster	Tiga Klaster	Empat Klaster
2	0.0790374	0.127169	0.141204
3	0.166808	0.284507	0.371297

Berdasarkan pengujian diatas, nilai yang terbaik adalah nilai dengan pembobotan 2. Maka, selanjutnya akan dilakukan pengujian terhadap pembentukan masing-masing klaster dengan inputan jumlah klaster yaitu dua klaster, tiga klaster, dan empat klaster dengan pembobotan 2, maksimal iterasi 1000, dan error yang ditoleransi 0,005. Hasil pengujian terhadap klaster dapat dilihat sebagai berikut :

1. Dua Klaster



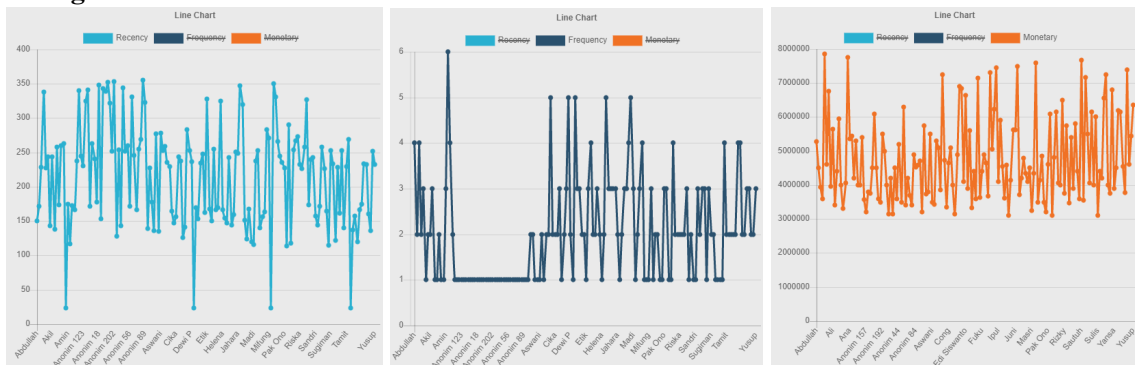
Gambar 2 Nilai RFM Cluster 1



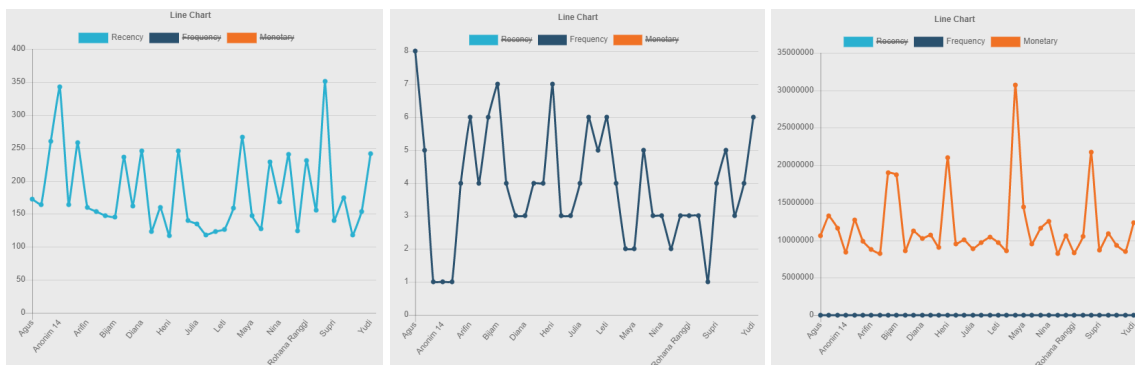
Gambar 3 Nilai RFM Cluster2

Kesimpulan dari gambar 2 dan 3 adalah hasil pada *cluster* satu (C1) menunjukkan ciri pelanggan yang tidak memiliki loyalitas terhadap toko. Ciri tersebut dapat dilihat pada nilai rata-rata RFM yang rendah yaitu frekuensi hanya 1, nilai *monetary* sekitar 2.024.213, serta nilai *recency* 247 dengan nilai minimal *monetary* 20.000, maksimal *monetary* 5.100.000 dan jumlah anggota 489 pelanggan sedangkan hasil pada *cluster* dua (C2) menunjukkan ciri pelanggan yang memiliki loyalitas tinggi terhadap toko. Ciri tersebut dapat dilihat pada nilai rata-rata RFM yang tinggi namun masih berfluktuatif yaitu frekuensi sebesar 3, nilai *monetary* sekitar 8.652.614, serta nilai *recency* 190 dengan nilai minimal *monetary* 5.100.000, maksimal *monetary* 30.750.000 dan jumlah anggota 88 pelanggan.

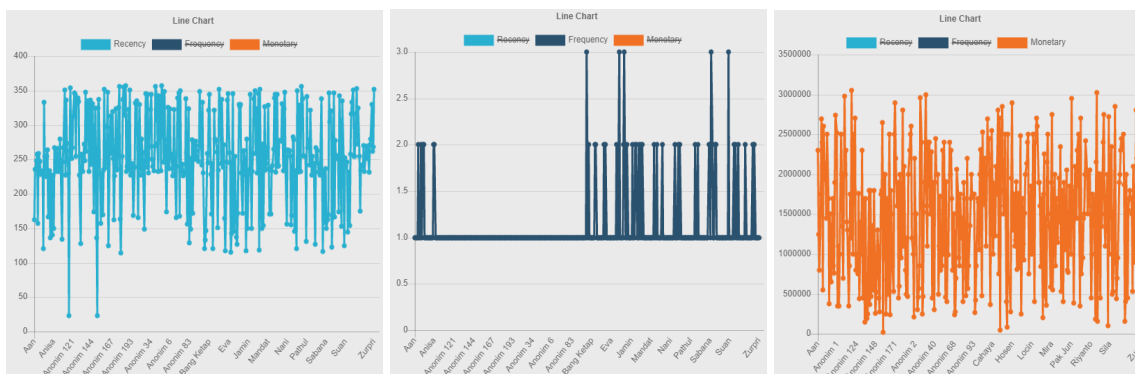
2. Tiga Klaster



Gambar 4 Nilai RFM Cluster1



Gambar 5 Nilai RFM Cluster 2

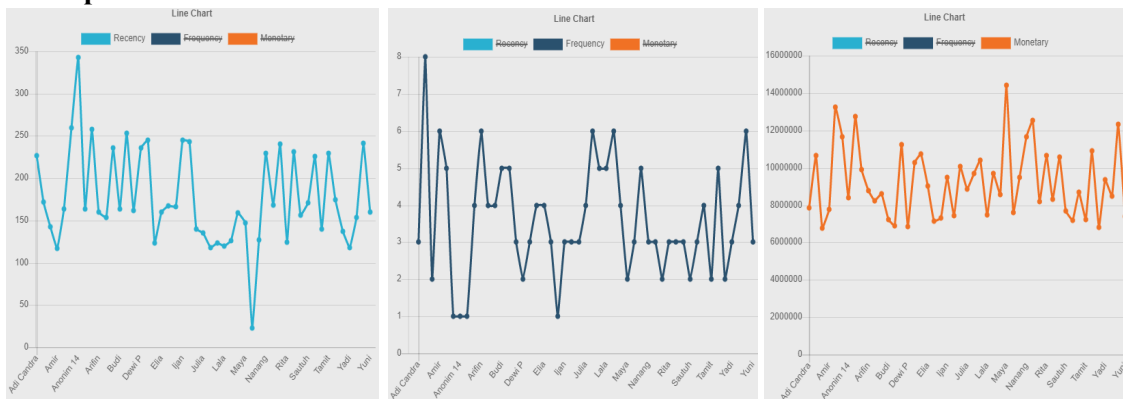


Gambar 6 Nilai RFM cluster 3

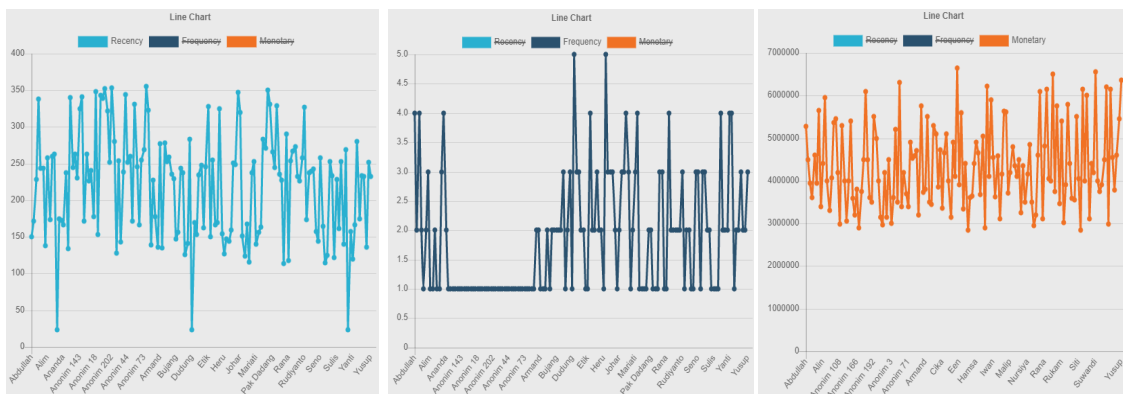
Kesimpulan dari gambar 4, 5, dan 6 adalah hasil pada *cluster* satu (C1) menunjukkan ciri pelanggan yang memiliki nilai loyalitas yang cukup tinggi terhadap toko. Ciri tersebut dapat

dilihat pada nilai rata-rata RFM yang cukup besar yaitu frekuensi sebesar 2, nilai *monetary* sekitar 4.753.052, serta nilai *recency* 215 dengan nilai minimal *monetary* 3.100.000, maksimal *monetary* 7.850.000 dan jumlah anggota 154 pelanggan. Hasil pada *cluster* dua (C2) menunjukkan ciri pelanggan yang memiliki loyalitas yang tinggi terhadap toko. Ciri tersebut dapat dilihat pada nilai rata-rata RFM yang tinggi yaitu frekuensi sebesar 4, nilai *monetary* sekitar 11.765.132, serta nilai *recency* 182 dengan nilai minimal *monetary* 8.210.000, maksimal *monetary* 30.750.000 dan jumlah anggota 38 pelanggan. Hasil pada *cluster* tiga (C3) menunjukkan ciri pelanggan yang tidak memiliki loyalitas terhadap toko. Ciri tersebut dapat dilihat pada nilai rata-rata RFM yang kecil yaitu frekuensi hanya 1, nilai *monetary* sekitar 1.486.299, serta nilai *recency* 253 dengan nilai minimal *monetary* 20.000, maksimal *monetary* 3.050.000 dan jumlah anggota 385 pelanggan.

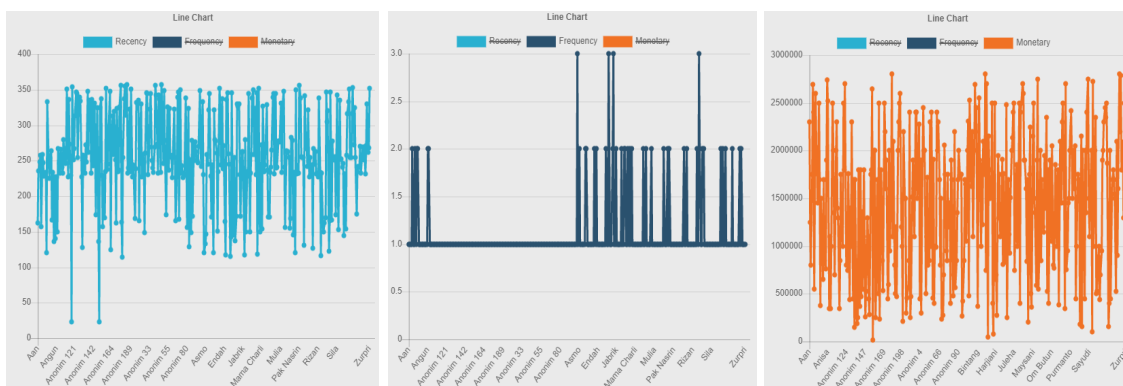
3. Empat Klaster



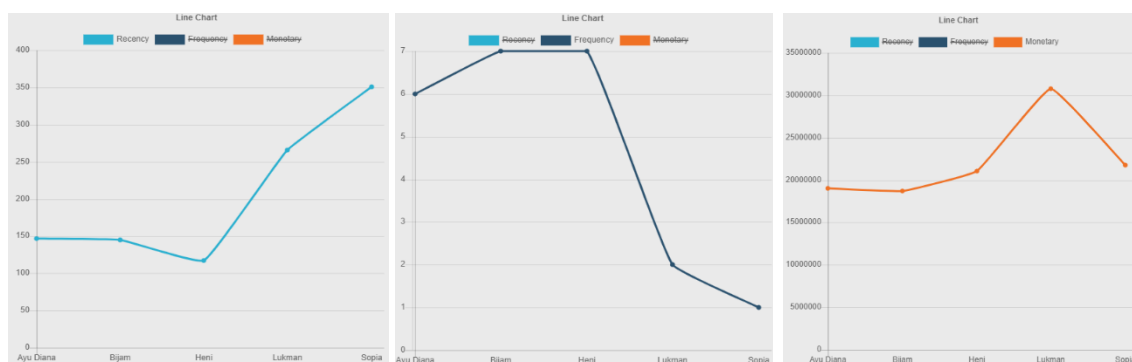
Gambar 7 Nilai RFM Cluster 1



Gambar 8 Nilai RFM Cluster 2



Gambar 9 Nilai RFM Cluster 3



Gambar 10 Nilai RFM Cluster 4

Kesimpulan dari gambar 7, 8, 9 dan 10 adalah hasil pada *cluster* satu (C1) menunjukkan ciri pelanggan yang memiliki nilai loyalitas yang cukup tinggi terhadap toko. Ciri tersebut dapat dilihat pada nilai rata-rata RFM yang baik namun masih berfluktuatif yaitu frekuensi sebesar 2, nilai *monetary* sekitar 4.347.248, serta nilai *recency* 220 dengan nilai minimal *monetary* 2.850.000, maksimal *monetary* 6.640.000 dan jumlah anggota 149 pelanggan. Hasil pada *cluster* dua (C2) dan empat (C4) menunjukkan ciri pelanggan yang memiliki loyalitas yang sama terhadap toko dengan jumlah anggota 49 pelanggan untuk C2 dan 5 pelanggan untuk C4. Hasil pada *cluster* tiga (C3) menunjukkan ciri pelanggan yang tidak memiliki loyalitas terhadap toko. Ciri tersebut dapat dilihat pada nilai rata-rata RFM yang kecil yaitu frekuensi hanya 1, nilai *monetary* sekitar 1.443.275, serta nilai *recency* 254 dengan nilai minimal *monetary* 20.000, maksimal *monetary* 2.800.000 dan jumlah anggota 374 pelanggan.

Berdasarkan semua *output* gambar *chart* pada proses pembentukan dua klaster, tiga klaster dan empat klaster didapat bahwa hasil terbaik didapat pada saat pembentukan tiga klaster. Hasil pada klaster ketiga menunjukkan nilai RFM yang tinggi, hal ini menunjukkan bahwa pelanggan pada klaster tersebut memiliki tingkat loyalitas yang tinggi terhadap toko dan dapat dijadikan sebagai bahan pertimbangan bagi toko untuk memberikan diskon terhadap pelanggan yang terdapat pada klaster tersebut.

4. KESIMPULAN

Kesimpulan yang dapat diperoleh berdasarkan hasil pengujian yang telah dilakukan terhadap sistem prediksi *customer loyalty* dengan metode RFM dan *Fuzzy C-Means* ini adalah sebagai berikut:

1. Semua fitur yang terdapat dalam modul-modul yang ada dapat dijalankan dan berfungsi dengan semestinya dan tentunya dapat digunakan dengan baik sesuai spesifikasi yang ada.
2. Hasil pengujian terhadap pembentukan klaster menunjukkan bahwa hasil terbaik terdapat pada pembentukan klaster dengan pembobotan 2 karena nilai PCI dan PEI yang didapat lebih baik dari pembobotan 3.
3. Hasil pengujian evaluasi *Partition Coefficient Index* dan *Partition Entropy Index* menunjukkan hasil pembentukan terbaik terdapat pada pembentukan dua klaster dengan pembobotan 2 didapatkan nilai PCI sebesar 0,892844 dan PEI sebesar 0,0792693.
4. Hasil pengujian evaluasi *Partition Coefficient Index* dan *Partition Entropy Index* dengan pembentukan tiga klaster juga menghasilkan nilai yang baik dengan nilai PCI sebesar 0,843572 dan PEI sebesar 0,126067 karena nilai PCI yang didapat mendekati angka 1 dan

nilai PEI yang didapat mendekati angka 0. Walaupun demikian, pembentukan dua klaster lebih baik dari tiga klaster.

5. Hasil pengelompokkan pembentukan dua klaster dengan pembobotan 2 menunjukkan bahwa pelanggan cukup sering berbelanja terhadap toko sampai Desember 2018 dan pembentukan tiga klaster dengan pembobotan 2 menunjukkan bahwa pelanggan sering melakukan transaksi serta nilai *monetary* yang cukup besar terhadap toko.

DAFTAR PUSTAKA

- [1] Cheng, Ching-Hsue, Chen, You-Shyang., 2009, Classifying the segmentation of customer value via RFM model, *Journal Expert System with Application: An International Journal*, Vol. 36, No. 3.
- [2] Mirkin, Boris., 2005, *Clustering for Data Mining: A Data Recovery Approach*, Chapman & Hall/CRC, London.
- [3] Tan, Pang-Ning, Steinbach, Michael, dan Kumar, Vipin, 2006, *Introduction to Data Mining*, Pearl International Edition, h.578.
- [4] Jaya, Tri Sandhika, Adi, Kusworo, Norita, Beta, Sistem Pemilihan Perumahan dengan Metode Kombinasi Fuzzy C-Means Clustering dan Simple Additive Weighting. https://www.researchgate.net/publication/304217293_Sistem_Pemilihan_Perumahan_dengan_Metode_Kombinasi_Fuzzy_CMeans_Clustering_dan_Simple_Additive_Weighting, diakses tanggal 02 September 2019.
- [5] Datanovia, Fuzzy C-Means Clustering Algorithm. <https://www.datanovia.com/en/lessons/fuzzy-clustering-essentials/fuzzy-c-means-clustering-algorithm>, diakses tanggal 08 September 2019.
- [6] Ruang Guru.co, Rumus Koefisien Variasi, <https://ruangguru.co/rumus-koefisien-variasi>, diakses tanggal 25 November 2019.
- [7] Setiawan, Adi, 2012, Perbandingan Koefisien Variasi Antara 2 Sampel Dengan Metode Bootstrap, *JdC*, Vol. 1, No 1, h.19-20.
- [8] Selviana, Nur Indah dan Mustakim., 2016, Analisis Perbandingan K-Means dan Fuzzy C-Means Untuk Pemetaan Motivasi Balajar Mahasiswa. *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI)* 8.
- [9] Hardiani, Tikaridha., 2018, Segmentasi Nasabah Simpanan Menggunakan Fuzzy C Means dan Fuzzy RFM (Recency, Frequency, Monetary) Pada BMT XYZ, *Jurnal Ilmiah NERO*, Vol. 3, No. 3.