



Segmentasi Loyalitas Pelanggan Berbasis RFM (Recency, Frequency, Monetary) Menggunakan K-Means pada PD. Persada Ikan

Yosia Oktavian Pailan¹, Yulison Herry Chrisnanto², Asep Id Hadianna³

¹²³ Jurusan Informatika, fakultas Sains dan Informatika, Universitas Jenderal Achmad Yani
yosiaokta91170@gmail.com

Abstract

Customer loyalty for the company is very important if the competition between similar companies is high enough which results in a threat to the company. Customer loyalty is very useful to determine the level of customer loyalty to the company. Customer segmentation is also needed to group customers who have the same characteristics into one so as to simplify the management process. The analysis used is the RFM (Recency, Frequency, Monetary) model to analyze customer buying behavior in terms of Recency (last transaction time span), Frequency (number of transactions), and Monetary (rupiah issued). The grouping method used is K-Means. The data used in this study are historical data on fish feed purchases from 2015 to 2017. The application of RFM analysis and the K-Means method resulted in 4 clusters based on Elbow calculations. The results in this study obtained the number of objects in cluster 1 as many as 142 customers, cluster 2 as many as 28 customers, cluster 3 as many as 41 customers and cluster 4 as many as 41 customers. The accuracy level of the cluster is measured using Silhouette Coefficient with results close to 1 which means the clustering is quite good. The interpretation of the RFM shows that 16.27% of customers have a high potential for loyalty, while 11.11% of customers have the potential as loyal customers, and the remaining 56.35% have a low level of loyalty. It can be concluded that this research can classify the level of customer loyalty using RFM analysis and the K-Means algorithm.

Keywords: RFM, k-means, loyalty, segmentation.

Abstrak

Loyalitas pelanggan bagi perusahaan sangatlah penting apabila persaingan antar perusahaan sejenis cukup tinggi yang berakibat sebagai ancaman bagi perusahaan. Loyalitas pelanggan sangat berguna untuk mengetahui tingkat kesetiaan pelanggan terhadap perusahaan. Segmentasi pelanggan juga diperlukan untuk mengelompokkan pelanggan yang memiliki kesamaan karakteristik menjadi satu sehingga mempermudah dalam proses pengelolaan. Analisis yang digunakan adalah dengan model RFM (Recency, Frequency, Monetary) untuk menganalisis perilaku pembelian pelanggan dari segi Recency (rentang waktu transaksi terakhir), Frequency (jumlah transaksi), dan Monetary (rupiah yang dikeluarkan). Metode pengelompokan yang digunakan adalah K-Means. Data yang digunakan dalam penelitian ini adalah data riwayat transaksi pembelian pakan ikan pada kurun waktu tahun 2015 hingga 2017. Penerapan analisis RFM dan metode K-Means menghasilkan 4 cluster berdasarkan perhitungan Elbow. Hasil dalam penelitian ini didapatkan jumlah objek pada cluster 1 sebanyak 142 pelanggan, cluster 2 sebanyak 28 pelanggan, cluster 3 sebanyak 41 pelanggan dan cluster 4 sebanyak 41 pelanggan. Tingkat akurasi cluster diukur menggunakan Silhouette Coefficient dengan hasil mendekati 1 yang berarti klusterisasi cukup baik. Interpretasi dari RFM menunjukkan 16,27% pelanggan memiliki potensi loyalitas cukup tinggi, sedangkan 11,11% pelanggan memiliki potensi sebagai pelanggan loyal, dan sisanya 56,35 % tingkat loyalitas rendah. Penelitian ini dapat disimpulkan dapat mengelompokkan tingkat loyalitas pelanggan menggunakan analisis RFM dan algoritma K-Means.

Kata kunci: RFM, k-means, loyalitas, segmentasi

1. Pendahuluan

Mengetahui loyalitas pelanggan bagi perusahaan sangat penting apabila ada persaingan antar perusahaan sejenis yang berakibat sebagai ancaman bagi perusahaan. Loyalitas pelanggan sangat berguna dalam membuat keputusan strategi *marketing* bagi perusahaan terhadap setiap pelanggan. Upaya untuk mengetahui loyalitas pelanggan bagi perusahaan yaitu dengan konsep segmentasi pelanggan. Segmentasi pelanggan adalah proses pembelajaran *unsupervised* yang mengklasifikasikan pelanggan ke dalam kelas yang telah

ditentukan dari kelompok pelanggan sebelumnya dari dalam kelas yang belum ditentukan. Informasi grup yang ditemukan sangat berguna dalam perumusan strategi promosi dan kebijakan penetapan harga untuk meningkatkan tingkat respons pelanggan serta meningkatkan laba [1]. Segmentasi pelanggan diperlukan untuk mengelompokkan pelanggan yang memiliki kesamaan karakteristik. Hal ini diperlukan untuk mengetahui perilaku (*behavior*) konsumen sehingga akan membantu dalam penerapan strategi pemasaran (*marketing*) yang tepat untuk meningkatkan

pendapatan perusahaan [2]. Selain segmentasi pelanggan, teknologi informasi sangat berpengaruh bagi kemajuan sebuah perusahaan. Teknologi informasi membantu memberikan layanan atau produk yang lebih baik kepada pelanggan [3]. Kebutuhan akan pakan ikan sangat diperlukan bagi pemilik kolam terapung di waduk Cirata. Jenis ikan yang sering dibudidayakan di kawasan waduk Cirata adalah ikan mas dan ikan nila. Produk pakan yang diperlukan pemilik kolam terapung beraneka ragam merk sesuai dengan komposisi yang dibutuhkan ikan untuk berkembang dalam waktu kurang lebih 3 sampai 4 bulan sudah bisa dipanen. Setiap tahun terdapat penambahan jumlah perusahaan yang muncul seiring menurunnya jumlah pelanggan yang dimiliki sehingga perusahaan perlu mengetahui tingkat kesetiaan pelanggannya.

Model Recency Frequency Monetary (RFM) adalah model berbasis perilaku digunakan untuk menganalisis perilaku pelanggan dan dibuatkan prediksi berdasarkan perilaku database. Model RFM ini merupakan metode yang sudah lama dan populer untuk mengukur hubungan dengan pelanggan [4]. Dalam penelitian terdahulu yang membahas clustering menggunakan K-Means clustering untuk melakukan segmentasi pelanggan. Alasan metode ini digunakan adalah karena metode ini merupakan metode interaktif yang mudah diinterpretasikan, diterapkan, dan bersifat dinamis pada data yang tersebar [5]. Dalam menentukan jumlah Cluster yang optimal bisa divalidasi dengan menggunakan Silhouette Index. Tujuan dari teknik validasi cluster adalah untuk mengevaluasi hasil clustering, hasil evaluasi ini dapat digunakan untuk menentukan jumlah cluster pada dataset. Salah satu teknik validasi cluster adalah Silhouette Index yaitu metode penafsiran untuk validasi klaster pada objek-objek. Teknik ini memberikan representasi grafis singkat tentang seberapa baik setiap objek terletak dalam clusternya [6]. Dalam penelitian yang dilakukan untuk mengelompokkan customer diperlukan suatu metode pengelompokan atau clustering. Saat ini telah banyak algoritma yang digunakan untuk segmentasi customer seperti metode SOM (Self Organizing Map), Fuzzy C- Means dan K-Means. Metode Algoritma K- Means adalah metode Clustering yang paling populer digunakan karena beberapa kelebihan yang dimiliki antara lain algoritma ini sederhana dan mudah diterapkan [7].

Penelitian ini menggunakan metode K-Means untuk mengelompokkan tingkat loyalitas pelanggan berdasarkan transaksi pembelian pakan ikan sehingga dapat dilakukan analisis RFM untuk mengetahui tingkat kesetiaan pelanggan bagi perusahaan serta bantuan perhitungan *silhouette* untuk menguji tingkat akurasi hasil pengelompokan loyalitas pelanggan. Untuk mengekstraksi data menjadi sebuah informasi akan dibantu oleh data mining.

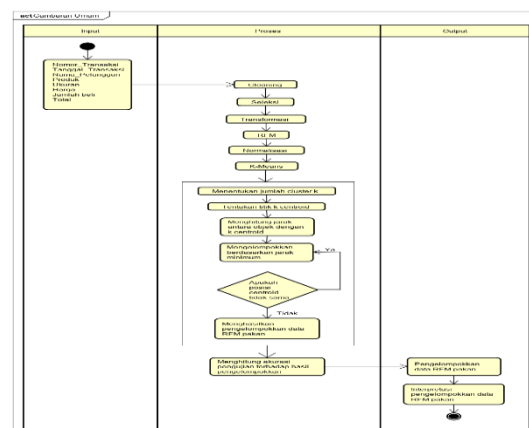
Berdasarkan latar belakang diatas, maka penulis dapat merumuskan permasalahan yaitu bagaimana cara

mengelompokkan tingkat kesetiaan pelanggan menggunakan sistem yang akan dibangun ?

Penelitian ini memiliki tujuan yaitu membangun sistem yang dapat mengelompokkan dan mengetahui tingkat loyalitas pelanggan berdasarkan transaksi pembelian pakan ikan melalui analisis interpretasi RFM.

2. Metode Penelitian

Dalam penelitian ini dilakukan beberapa tahap dalam mencapai tujuan dari pembuatan sistem pengelompokan penjualan untuk mengetahui tingkat loyalitas pelanggan yaitu dimulai dari tahap persiapan data utama, tahap analisis RFM, tahap segmentasi menggunakan K-Means Clustering, tahap visualisasi hasil segmentasi, tahap pengujian dan evaluasi. Gambaran umum penelitian dapat dilihat pada Gambar 1.



Gambar 1 Metode Penelitian

2.1. Data Masukan

Data masuk adalah data transaksi dalam kurun 3 tahun dimulai tahun 2015 sampai dengan 2017 dengan jumlah transaksi sebanyak 9007 record.

2.2. Cleaning

Pada tahap ini dilakukan pembersihan terhadap data yang akan digunakan sebagai bahan dalam pembuatan sistem pengelompokan transaksi. Pada tahap ini pembersihan yang dilakukan terhadap atribut – atribut data yang terdapat pada data transaksi yang tidak diperlukan atau tidak digunakan akan dihilangkan.

2.3. Seleksi

Seleksi data transaksi pakan ikan bertujuan untuk menyeleksi data transaksi yang akan digunakan dalam penelitian. Pada tahap ini data yang dihilangkan adalah data nama pakan dan ukuran pakan sehingga data yang digunakan adalah nama pelanggan, tanggal transaksi, jumlah beli dan total pembelian.

2.4. RFM

Sebelum proses segmentasi dengan clusterisasi, maka perlu ditetapkan atribut-atribut sebagai nilai R, F dan M pada dataset yang ada. Perhitungan nilai Recency (R)

mengacu pada interval antara selisih waktu dengan tanggal transaksi pada setiap data penjualan. Sedangkan untuk nilai Frequency (F) mengacu pada jumlah transaksi dalam periode tertentu dan nilai Monetary (M) mengacu pada jumlah perkalian kuantitas barang yang dibeli dengan harga satuan sesuai dengan harga barang [1].

2.5. Normalisasi

Hasil nilai RFM yang sudah ditentukan akan dilakukan tahap normalisasi dari setiap atribut RFM dengan menggunakan persamaan :

$$v' = \frac{v - \min_a}{\max_a - \min_a} (\text{newmax}_a - \text{newmin}_a) + \text{newmin}_a \quad (1)$$

Dengan v' adalah nilai setelah dinormalisasi, v adalah nilai sebelum dinormalisasi, \min_a adalah nilai minimal setiap variable, \max_a adalah nilai maksimal setiap variable, newmax_a adalah rentan maksimal x dengan nilai 1 dan newmin_a adalah rentan nilai x dengan nilai 0.

2.6. Clustering K-Means

Clustering adalah teknik pembelajaran unsupervised dan dapat didefinisikan sebagai pembagian populasi heterogen menjadi kelompok yang lebih homogen. Terdapat berbagai aplikasi yang menggunakan teknik clustering yaitu segmentasi pelanggan, segmentasi gambar, pengambilan informasi, pengelompokan halaman web dan ilmiah serta teknik analisis [8].

K-Means merupakan metode yang digunakan dalam algoritma clustering berbasis jarak yang membagi data ke dalam sejumlah cluster dan algoritma ini hanya bekerja pada atribut numeric [9]

Proses segmentasi pelanggan dilakukan melalui tahapan penentuan nilai k sebagai representasi dari banyaknya cluster yang akan dibentuk. Untuk menentukan jumlah nilai k ideal dari dataset menggunakan metode elbow.

Perhitungan jarak kedekatan objek dengan centroid dapat menggunakan perhitungan jarak Euclidean Distance, yaitu persamaan untuk menghitung jarak antara objek data melalui perhitungan akar perbedaan objek data yang berpasangan. Jarak tersebut digunakan untuk menentukan cluster dari anggota pada proses K-Means Clustering. Berikut adalah persamaan Euclidean Distance.

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (2)$$

Dengan D_e adalah Euclidian Distance, i adalah banyaknya data, (x,y) adalah koordinat data dan (a,b) adalah koordinat centroid.

2.7. Silhouette Coefficient

Silhouette Coefficient merupakan persamaan yang digunakan untuk mengukur seberapa akurat dari

pengelompokan yang sudah dilakukan sebelumnya. Persamaan ini memberikan nilai untuk keakuratan pada hasil pengelompokan dengan rentan nilai 0 – 1 jika nilai sudah mendekati angka satu maka hasil dari pengelompokan semakin baik [10]

Untuk menguji tingkat akurasi kluster, maka dilakukan pengujian menggunakan metoda Silhouette sesuai dengan persamaan :

$$S_i = \frac{b_i - a_i}{\text{Max}\{a_i, b_i\}} \quad (3)$$

Dengan S_i adalah silhouette index, a_i adalah rata-rata jarak antara titik i dengan seluruh titik dalam A (cluster dimana titik i berada) dan b_i adalah rata-rata jarak antara titik i ke seluruh titik dalam cluster selain A.

2.8. Tahap Pengujian dan Evaluasi

Dalam tahap ini dilakukan pada setiap fungsi atau prosedur yang terdapat di dalam modul, hal ini dilakukan untuk mencegah kesalahan-kesalahan pada sistem yang telah dibuat.

3. Hasil dan Pembahasan

Berikut data transaksi yang telah diubah dalam analisis RFM pada Tabel 1.

Tabel 1. Data Transaksi diubah dalam analisis RFM

No	Tanggal	Nama	Jumlah Beli	Total
1	01/01/2015	Bayu Suheryanto	13	Rp 3.185.000
2	01/01/2015	Bayu Suheryanto	12	Rp 3.900.000
3	01/01/2015	Deden Saepudin	9	Rp 3.465.000
4	01/01/2015	Deden Saepudin	13	Rp 3.185.000
5	01/01/2015	Endang Supriatna	15	Rp 5.400.000
6	01/01/2015	Endang Supriatna	17	Rp 6.120.000
7	01/01/2015	Nurdiansyah	9	Rp 3.240.000
8	01/01/2015	Nurdiansyah	11	Rp 3.300.000
9	01/01/2015	Rudi Suci Firdaus	18	Rp 6.480.000
10	01/01/2015	Suhalmi Supriatna	15	Rp 3.900.000
....
9005	30/12/2017	Lalan Permana	13	Rp 3.900.000

Metode normalisasi melakukan transformasi atribut numerik dalam skala yang lebih kecil yaitu dengan batas terendah 0 dan batas tertinggi 1. Berikut adalah perhitungan normalisasi Min - Max seperti dalam Tabel 2 dan 3.

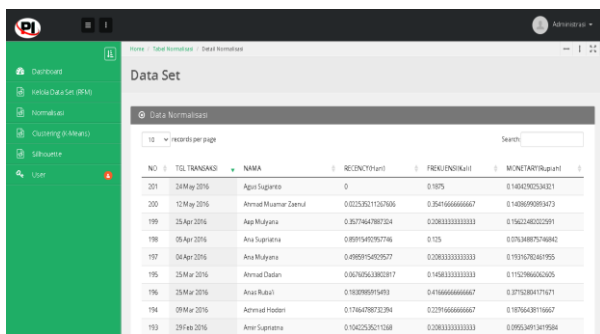
Tabel 2. Data RFM transaksi

	Recency	Frequency	Monetary
C-001	140	4	12280000
C-002	164	6	36380000
C-003	45	6	25325000
C-004	53	2	6945000
C-005	144	4	13800000
C-006	37	2	6705000
C-007	132	3	12480000
C-008	176	5	24910000
C-009	120	11	71190000
C-010	44	3	22380000

Tabel 3. Data RFM transaksi hasil normalisasi

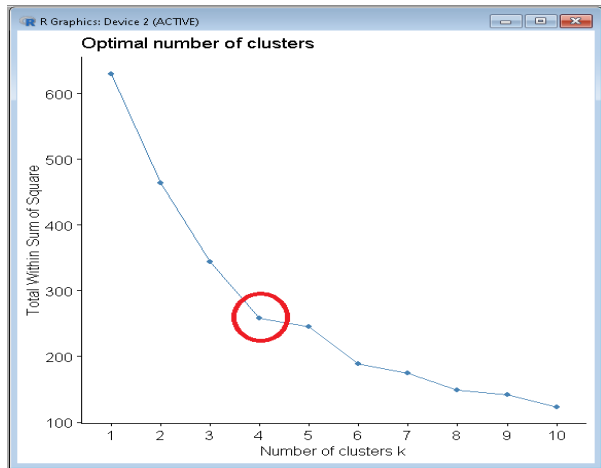
	nRecency	nFrequency	nMonetary
C-001	0,74	0,22	0,09
C-002	0,91	0,44	0,46
C-003	0,06	0,44	0,29
C-004	0,12	0,00	0,00
C-005	0,77	0,22	0,11
C-006	0,00	0,00	0,00
C-007	0,68	0,11	0,09
C-008	0,00	0,33	0,28
C-009	0,60	1,00	1,00
C-010	0,05	0,11	0,24

Berikut program hasil perhitungan normalisasi min-max dalam Gambar 2.



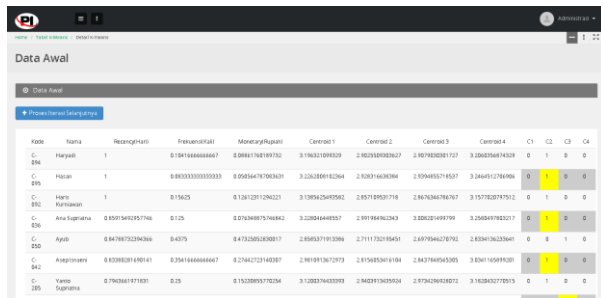
Gambar 2 Hasil perhitungan normalisasi

Objek k dipilih secara acak sebagai dasar pusat cluster (centroid awal) sebanyak k (jumlah cluster). Pada penentuan jumlah k telah ditentukan sebanyak 4 menurut hasil pengujian dalam program R-Language pada Gambar 3.



Gambar 3 Hasil pengujian penentuan jumlah cluster optimal

Berikut program hasil perhitungan clustering k-means dalam Gambar 4.



Gambar 4 Hasil perhitungan clustering K-Means

Dalam analisis hasil RFM diperoleh dari hasil interpretasi RFM penjualan pakan ikan yang telah dilakukan. Berikut hasil interpretasi RFM dalam tabel 4.

Tabel 4. Interpretasi RFM

No	Cluster	Jumlah Pelanggan	R	F	M
1	Cluster 1	142	2	2	2
2	Cluster 2	28	2	4	4
3	Cluster 3	41	5	5	4
4	Cluster 4	41	5	4	3

Setelah didapatkan nilai dari masing-masing atribut Recency, Frequency dan monetary selanjutnya yang perlu dilakukan adalah memberikan score kuantitatif untuk masing-masing nilai tersebut agar bisa dilakukan pemrosesan di tahap selanjutnya [11]. Pemberian nilai kuantitatif ini digunakan dengan menggunakan rentang skala 1-5 yang memiliki rincian seperti yang ditentukan pada Tabel 5.

Tabel 5. Penentuan Nilai Atribut RFM

No	Recency	Frequency	Monetary
1	13-85	27-34	15000000 - 20000000
2	86-156	19-26	50000000 - 150000000
3	157-215	11-18	25000000 - 50000000
4	216-277	3-10	2000000 - 25000000
5	278-373	0-1	< 2000000

Berdasarkan skala score yang telah dibuat, maka hasil evaluasi atribut RFM dari hasil interpretasi program menggunakan algoritma K-means ditampilkan seperti pada Tabel 6.

Tabel 6. Hasil klasterisasi dan rata-rata score RFM

Cluster	Jumlah Data Setiap Cluster	Analisis RFM
1	142	↓R↓F↓M
2	28	↓R↑F↑M
3	41	↑R↑F↑M
4	41	↑R↑F↑M

Berdasarkan hasil scoring serta hasil klasterisasi yang sudah dilakukan melalui perangkat lunak yang dikembangkan, maka dapat diinterpretasikan karakteristik pelanggan berdasarkan atribut Recency, Frequency dan Monetary seperti yang disajikan pada Tabel 7.

Tabel 7. Interpretasi hasil cluster RFM

No	Interpretasi
1	Nilai R (Recency) dengan skor 2, artinya pelanggan pada kelompok ini sudah lama tidak melakukan transaksi. Demikian juga pelanggan pada kelompok ini memiliki F (Frequency) yang rendah dengan skor 2, itu artinya pelanggan pada kelompok ini jarang melakukan transaksi pembelian pada kurun waktu tertentu serta skor M (Monetary) rendah dengan skor 2, di mana pelanggan pada kelompok ini kurang memberikan nilai finansial bagi perusahaan.
2	Nilai R (Recency) dengan skor 2, artinya pelanggan pada kelompok ini sudah lama tidak melakukan transaksi. Namun demikian pelanggan pada kelompok ini memiliki F (Frequency) yang baik dengan skor 4, itu artinya pelanggan pada kelompok ini sering melakukan transaksi pembelian pada kurun waktu tertentu. Sedangkan skor M (Monetary) sangat baik dengan skor 4, di mana pelanggan pada kelompok ini memberikan nilai finansial tinggi bagi perusahaan.
3	Nilai R (recency) dengan skor 5, artinya waktu transaksi terakhir dilakukan belum lama, hal tersebut memberikan potensi pelanggan tersebut untuk melakukan transaksi kembali. Demikian juga dengan pelanggan pada kelompok ini memiliki F (Frequency) yang tinggi dengan skor 5, itu artinya pelanggan pada kelompok ini sering melakukan transaksi pembelian pada kurun waktu tertentu. Sedangkan skor M (Monetary) sangat baik dengan skor 4, di mana pelanggan pada kelompok ini memberikan nilai finansial tinggi bagi perusahaan.
4	Nilai R (recency) dengan skor 5, artinya waktu transaksi terakhir dilakukan belum lama, hal tersebut memberikan potensi pelanggan tersebut untuk melakukan transaksi kembali. Demikian juga pelanggan pada kelompok ini memiliki F (Frequency) yang baik dengan skor 4, itu artinya pelanggan pada kelompok ini sering melakukan transaksi pembelian pada kurun waktu tertentu. Sedangkan skor M (Monetary) cukup baik dengan skor 3, di mana pelanggan pada kelompok ini memberikan nilai finansial bagi perusahaan.

4. Kesimpulan

Hasil analisis interpretasi RFM yang dilakukan memiliki objek yang terbentuk sebanyak 4 cluster dengan 142 pelanggan berada di cluster 1 dengan indikasi loyalitas rendah, 28 pelanggan berada di cluster 2 dengan indikasi loyalitas sedang, 41 pelanggan berada di cluster 3 dengan indikasi loyalitas tinggi dan karakteristik 41 pelanggan berada di cluster 4 dengan indikasi loyalitas sedang sehingga terdapat 142 pelanggan berpotensi tidak loyal pada P.D Persada Ikan. Adapun pengujian akurasi dalam sistem terdapat data hasil dari pengelompokan transaksi penjualan pakan ikan selama 3 tahun, hasil keakurasian data mencapai 0.770 mendekati angka 1. Saran untuk sistem ini adalah dengan mencoba menggunakan metode lain untuk mengukur dan mendapatkan hasil lebih baik.

Daftar Rujukan

- [1] H. C. Chang and H. P. Tsai, "Group RFM analysis as a novel framework to discover better customer consumption behavior," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 14499–14513, 2011.
- [2] B. E. Adiana, I. Soesanti, and A. E. Permanasari, "Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering," *J. Terap. Teknol. Inf.*, vol. 2, no. 1, pp. 23–32, 2018.
- [3] I. Soesanti, "Web-Based Monitoring System On The Production Process Of Yogyakarta Batik Industry," *J. Theor. Appl. Inf. Technol.*, vol. 87, no. 1, pp. 146–152, 2016.
- [4] F. Hadi, M. Mustakim, D. O. Rahmadia, F. H. Nugraha, N. P. Bulan, and S. Monalisa, "Penerapan K-Means Clustering Berdasarkan RFM Mofek Sebagai Pemetaan dan Pendukung Strategi Pengelolaan Pelanggan (Studi Kasus: PT. Herbal Penawar Alwahidah Indonesia Pekanbaru)," *J. Sains dan Teknol. Ind.*, vol. 15, no. 1, pp. 69–76, 2017.
- [5] A. D. Savitri, F. A. Bachtiar, and N. Y. Setiawan, "Segmentasi Pelanggan Menggunakan Metode K-Means Clustering Berdasarkan Model RFM Pada Klinik Kecantikan (Studi Kasus : Belle Crown Malang)," vol. 2, no. 9, pp. 2957–2966, 2018.
- [6] S. Monalisa, "Segmentasi Perilaku Pembelian Pelanggan Berdasarkan Model RFM dengan Metode K-Means," vol. 2, no. 1, pp. 9–15, 2018.
- [7] D. Yanto and Jamal, "Analisis RFM dan Algoritma K-Means untuk Clustering Loyalitas Customer," vol. 9, no. 1, pp. 0–8, 2019.
- [8] H. Güçdemir and H. Selim, "Integrating multi-criteria decision making and clustering for business customer segmentation," *Ind. Manag. Data Syst.*, vol. 115, no. 6, pp. 625–645, 2015.
- [9] D. S. Wardiani and N. Merlina, "Implementasi Data Mining untuk Mengetahui Manfaat RPTRA menggunakan Metode K-means Clustering," *PILAR*, vol. 15, no. 1, pp. 125–132, 2019.
- [10] M. Anggara, H. Sujiani, and N. Helfi, "Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokan Member Di Alvaro Fitness," *J. Sist. dan Teknol. Inf.*, vol. 1, no. 1, pp. 1–6, 2016.
- [11] Y. H. Chrisnanto and A. Kaniaingsih, "Pengelompokan Ekuitas Pelanggan Berbasis Recency Frequency Monetary (RFM) Menggunakan K-Means Clustering," vol. 2019, no. Sentika, pp. 13–14, 2019.