

Perancangan Aplikasi Segmentasi Pelanggan Menggunakan *K-Means* Dengan Model RFM

Nico Chandra¹, Indah Fenriana²

Fakultas Sains dan Teknologi, Universitas Buddhi Dharma

Jalan Imam Bonjol No. 41, Tangerang, Indonesia

Email: ¹Nicochan295@gmail.com, ²Indah.f88@gmail.com

Abstrak

Dalam dunia bisnis yang penuh persaingan, pelanggan menjadi aset berharga bagi setiap perusahaan. Keberhasilan sebuah perusahaan tidak hanya ditentukan oleh produk atau layanan yang berkualitas, tetapi juga oleh bagaimana perusahaan dapat memahami dan merespons kebutuhan serta preferensi pelanggan dengan baik. Perusahaan yang mampu menjalin hubungan yang kuat dengan pelanggan akan lebih mampu bertahan dan tumbuh dalam pasar yang kompetitif. Oleh karena itu, segmentasi pelanggan menjadi langkah strategis bagi perusahaan. Penelitian ini bertujuan untuk menerapkan metode *k-means* dalam segmentasi pelanggan dengan model RFM (*Recency*, *Frequency* dan *Monetary*). *K-Means* merupakan metode dalam machine learning yang digunakan untuk melakukan klustering atau pengelompokan data. Tujuan utama dari *K-Means* adalah untuk mengelompokkan data menjadi beberapa kelompok atau klaster berdasarkan kesamaan atribut atau fitur tertentu. Data yang digunakan pada penelitian ini menggunakan data transaksi penjualan sebanyak 3821, hasil implementasi menghasilkan 3 klaster pelanggan dengan jumlah 32 pelanggan yaitu 4 pelanggan dikategori sangat potensial, 19 pelanggan dikategori pelanggan potensial, 9 pelanggan dikategori tidak potensial. Dengan adanya aplikasi segmentasi pelanggan ini diharapkan dapat membantu perusahaan dalam menganalisis segmentasi pelanggan, untuk menyusun strategi pemasaran yang lebih tepat sasaran dan efektif. Sehingga dapat menciptakan hubungan yang kuat dengan pelanggan, meningkatkan loyalitas, dan mencapai kesuksesan dalam persaingan bisnis yang ketat.

Kata Kunci

Data mining, kmeans, Clustering, RFM, segmentation, customer

Latar Belakang

Dalam ranah bisnis, persaingan menjadi sesuatu yang tak terhindarkan, Seiring dengan kemajuan teknologi dan proses globalisasi, persaingan semakin meningkat di berbagai sektor bisnis. Hal ini disebabkan oleh kebutuhan untuk mencapai profitabilitas dan menjaga kelangsungan usaha. Persaingan dalam dunia bisnis bisa berasal dari mana saja, seperti hadirnya perusahaan-perusahaan baru di pasar, perubahan tren pasar, dan perkembangan teknologi. Persaingan juga bisa timbul dari kompetitor langsung maupun tidak langsung. Industri peternakan ayam broiler merupakan sektor usaha yang menghadapi persaingan ketat, permintaan tinggi akan daging ayam broiler, dan peluang keuntungan yang menjanjikan menjadi faktor yang mempengaruhi. PT. Satwa Prima Utama adalah perusahaan yang bergerak di sektor industri peternakan ayam yang didirikan pada tahun 2014. [12] Dalam menghadapi persaingan di industri ini, perusahaan menyadari pentingnya sebuah strategi dalam mengelola para pelanggannya dikarenakan pelanggan merupakan aset penting yang dapat membantu perusahaan berkembang serta bertahan dalam menghadapi persaingan usaha. Oleh karena itu, perusahaan berkomitmen untuk menerapkan segmentasi pelanggan sebagai upaya untuk

mengoptimalkan pemasaran dan pelayanan sesuai dengan kebutuhan masing-masing segmen pelanggan. [1] Segmentasi pelanggan adalah strategi yang diterapkan untuk mengelompokkan pelanggan ke dalam kategori-kategori yang berbeda berdasarkan perbedaan karakteristik, perilaku, atau kebutuhan mereka. Pendekatan ini memiliki kepentingan yang sangat penting karena memungkinkan perusahaan untuk menggunakan data pendukung yang bermanfaat dalam mengidentifikasi tingkat kesetiaan pelanggan, serta mengembangkan strategi yang efektif dan efisien dalam mencapai tujuan bisnis. [2] Dalam prosesnya, perusahaan menghadapi tantangan dalam melakukan segmentasi pelanggan karena jumlah data transaksi yang besar, sehingga mengelolanya secara manual menjadi tidak mungkin karena keterbatasan manusia dalam mengolah data. [3] Oleh karena itu, penerapan *data mining* menjadi sangat berguna untuk mengelola data dalam skala besar. Data mining merupakan serangkaian proses yang digunakan untuk menemukan informasi penting dari sekumpulan data besar dan kompleks yang sulit diolah secara manual. [10] *K-Means Clustering* dengan model RFM (*Recency, Frequency, Monetary*) merupakan pendekatan yang efektif dalam segmentasi pelanggan dalam analisis data bisnis. Model RFM memungkinkan perusahaan untuk mempertimbangkan tiga aspek utama dalam perilaku pelanggan, yaitu seberapa baru mereka bertransaksi (*Recency*), seberapa sering mereka berinteraksi dengan perusahaan (*Frequency*), dan total nilai transaksi atau pendapatan yang dihasilkan (*Monetary*). Dengan memanfaatkan model RFM, data pelanggan dapat dianalisis dan dikelompokkan ke dalam kelompok-kelompok berdasarkan kesamaan karakteristik menggunakan algoritma *K-Means*. [11] Pendekatan ini membantu perusahaan memahami pola pembelian dan kebutuhan pelanggan dengan lebih baik, dan memungkinkan perusahaan untuk mengambil tindakan pemasaran yang tepat sasaran dan mengoptimalkan strategi penjualan untuk meningkatkan kepuasan juga loyalitas pelanggan. Penelitian terkait dahulu yang telah dilakukan lainnya terkait dengan topik yang diteliti ini adalah sebagai berikut: [13] Atik Febriani¹, Syahfara Ashari Putri (2020) dengan judul “ Segmentasi Konsumen Berdasarkan Model *Recency, Frequency, Monetary* dengan Metode *K-Means* Dari hasil pengolahan data, terbentuk 3 kluster yang terbentuk dari 25 data pelanggan. Berdasarkan kluster yang terbentuk, dapat disimpulkan bahwa bahwa perilaku pembelian pelanggan memiliki pola yang berbeda-beda berdasarkan model RFM, pelanggan potensial didapatkan pada kluster 1 dan kluster 2 merupakan pelanggan kurang potensial.

Metode Penelitian

Data Mining

[4] *Data mining* ialah suatu ilmu pengetahuan yang dimanfaatkan guna menangani permasalahan dengan cara mengekstraksi informasi dari basis data yang berskala besar dengan menggabungkan beragam teknik berasal dari statistik, pembelajaran mesin, visualisasi data, identifikasi pola, serta manajemen basis data.

[5] *Data mining* adalah serangkaian proses yang berfungsi untuk mengeksplorasi dan mencari nilai berupa informasi serta relasi kompleks yang selama ini tersimpan dalam suatu basis data.

Clustering

[14] Clustering adalah teknik pengelompokan data dari sekumpulan data menurut karakteristik tertentu yang dimana nilai label belum diketahui.

Model *Recency, Frequency, Monetary* (RFM)

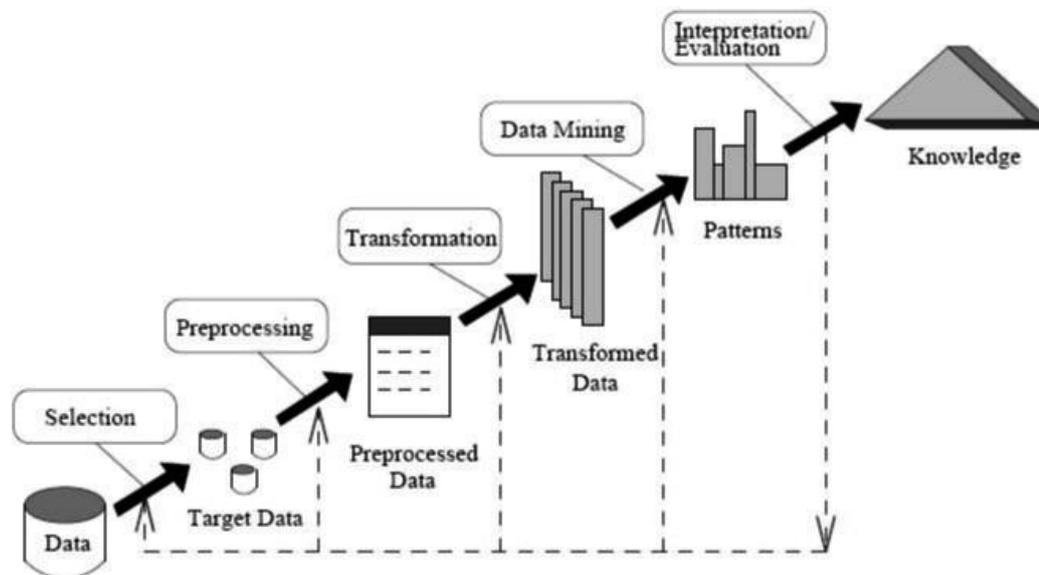
Model *Recency, Frequency, Monetary* pertama kali diperkenalkan oleh sebuah perusahaan konsultan bernama BAIN & Company pada tahun 1997. Model ini diciptakan sebagai alat untuk membantu perusahaan dalam menganalisis data pelanggan dan mengidentifikasi peluang pemasaran yang lebih baik berdasarkan pola pembelian pelanggan Model untuk mengkategorikan pelanggan melibatkan tiga faktor acuan, yaitu *Recency*, *Frequency*, dan

Monetary. Model RFM telah diterapkan secara luas dalam berbagai bidang, terutama dalam ranah pemasaran. Melalui integrasi model RFM, para pengambil keputusan dapat dengan cermat mengidentifikasi pelanggan signifikan dan memanfaatkan data ini untuk merumuskan pendekatan pemasaran yang sukses [2]. [16] Model ini memiliki tiga kriteria penting yaitu :

- a. *Recency* : Mengukur selisih antara tanggal analisis dan tanggal pembelian terakhir oleh pelanggan.
- b. *Frequency* : Menghitung total jumlah pembelian yang dilakukan oleh pelanggan dalam suatu periode.
- c. *Monetary* : Menyajikan total keseluruhan nominal pembelian yang dilakukan oleh pelanggan pada saat analisis.

Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database merupakan proses terstruktur dalam menganalisis data untuk memperoleh informasi yang tepat, baru, bermanfaat, dan mengungkapkan pola dari dataset yang kompleks. Langkah-langkah dalam proses ini meliputi seleksi dan pra-pemrosesan data, penerapan teknik seperti penambahan data, evaluasi pola, dan penafsiran hasil. KDD berperan dalam mengidentifikasi keterkaitan tersembunyi dan pola menarik dari data, yang dapat mendukung pengambilan keputusan yang lebih optimal. Hal ini dilakukan dengan mengekstraksi pengetahuan dari data yang sudah ada [6].



Gambar 1 Proses Knowledge Discovery in Database (KDD)

Sumber : [7]

Algoritma K-means

[8] Algoritma *K-means* salah satu metode *data mining* yang digunakan adalah untuk mengidentifikasi kelompok alami dari sebuah kasus berdasarkan pengelompokan data yang memiliki kemiripan. Dengan melakukan pengelompokan ini, dapat dilakukan analisis lebih lanjut terhadap data tersebut. Algoritma *K-Means* merupakan teknik dalam *data mining* yang mampu membentuk klaster dengan mengelompokkan data ke dalam satu atau lebih klaster. Hal ini memungkinkan data yang memiliki karakteristik serupa ditempatkan dalam satu klaster yang sama, sedangkan data yang menampilkan atribut yang berbeda ditempatkan dalam klaster

yang terpisah. Beroperasi secara eksklusif pada atribut numerik. [15] Berikut adalah langkah utama dari algoritma *K-Means Clustering* :

1. Inisialisasi *Centroid*

Langkah pertama adalah memilih secara acak K centroid awal sebagai pusat dari masing-masing kluster. K adalah jumlah kluster yang telah ditentukan sebelumnya.

2. Pengelompokan

Setiap data akan diberikan label kluster berdasarkan jarak Euclidean atau jarak lainnya antara data tersebut dengan centroid kluster terdekat. Data akan di-assign ke kluster yang memiliki centroid paling dekat.

3. Perbarui *Centroid*

Setelah pengelompokan awal, *centroid* dari masing-masing kluster akan diperbarui dengan menghitung rata-rata dari semua data yang termasuk dalam kluster tersebut.

4. Iterasi

Langkah 2 dan 3 akan diulang secara berulang-ulang hingga tercapai *konvergensi* atau batas iterasi tertentu. Iterasi akan berhenti ketika tidak ada perubahan dalam pengelompokan atau jika telah mencapai batas iterasi yang ditentukan.

5. Hasil Akhir

Setelah iterasi berakhir, kluster yang terbentuk akan berada pada posisi centroid yang stabil. Data telah sukses dikelompokkan menjadi K kluster berdasarkan kesamaan atributnya.

Min Max Normalization

Data yang terdapat dalam sebuah dataset seringkali menunjukkan variasi dalam rentang nilai. Kondisi ini dapat mempengaruhi hasil analisis data, sehingga memunculkan kebutuhan akan teknik normalisasi data untuk menyelaraskan rentang-nilai ini.[9] Normalisasi data melibatkan prosedur mengubah skala nilai atribut menjadi rentang yang lebih kecil sambil tetap mempertahankan bobot yang setara. Normalisasi *Min-Max* merupakan teknik yang melakukan konversi linear dengan menggunakan nilai minimum dan maksimum, menghasilkan distribusi data yang merata di seluruh rentang yang sama. Pendekatan ini mencapai konvergensi dengan kecepatan paling cepat dibandingkan dengan metodologi alternatif. Normalisasi *Min-Max* dapat dihitung melalui rumus berikut:

$$x_{New} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}}$$

Hasil

Pada tahap ini dilakukan proses KDD untuk mengolah data transaksi PT. Satwa Prima Utama dalam rentang waktu Januari 2021 – Juni 2023, dengan tahapan sebagai berikut :

Data Selection

Pada tahap ini, dilakukan proses persiapan dalam memilih data yang relevan dari kumpulan data transaksi yang besar agar dapat diterapkan dalam penelitian. Tujuan dari pemilihan data

adalah untuk mengurangi ukuran dataset sambil tetap mempertahankan data yang informatif dan penting. Data yang dipilih termasuk data tanggal transaksi penjualan dan total penjualan, yang akan digunakan untuk menentukan nilai *Recency*, *Frequency*, dan *Monetary* dalam penerapan model RFM. Berikut ini adalah data yang dipilih yang akan ditampilkan dalam tabel di bawah ini.

Tabel 1 *Data Selection*

| NO | TANGGAL | KODE | NAME | EKOR | KG | JUAK | BW | TOTAL |
|------|------------|-------|---------------------|-------|-------|--------|-------|------------|
| 1 | 02/01/2021 | 16 | MUHAMAD LUTFI | 1298 | 1400 | 19,000 | 1.08 | 26,600,000 |
| 2 | 02/01/2021 | 16 | MUHAMAD LUTFI | 759 | 1100 | 19,000 | 1.45 | 20,900,000 |
| 3 | 02/01/2021 | 16 | MUHAMAD LUTFI | 830 | 1200 | 19,000 | 1.45 | 22,800,000 |
| 4 | 02/01/2021 | 149 | MALHAN ADY TAMA | 552 | 800 | 19,000 | 1.45 | 15,200,000 |
| 5 | 02/01/2021 | 160 | SUJAJA | 394 | 450 | 19,000 | 1.14 | 8,550,000 |
| 6 | 02/01/2021 | 50 | YASIN | 1102 | 1300 | 19,500 | 1.18 | 25,350,000 |
| 7 | 02/01/2021 | 50 | YASIN | 1477 | 1600 | 19,500 | 1.08 | 31,200,000 |
| 8 | 02/01/2021 | 50 | YASIN | 998 | 980 | 19,500 | 0.98 | 19,110,000 |
| 9 | 02/01/2021 | 50 | YASIN | 677 | 501 | 14,500 | 0.74 | 7,264,500 |
| 10 | 03/01/2021 | 50 | YASIN | 876 | 1300 | 19,000 | 1.48 | 24,700,000 |
| 11 | 03/01/2021 | 50 | YASIN | 902 | 1300 | 19,000 | 1.44 | 24,700,000 |
| 12 | 03/01/2021 | 50 | YASIN | 740 | 1100 | 19,000 | 1.49 | 20,900,000 |
| 13 | 03/01/2021 | 50 | YASIN | 620 | 922.5 | 19,000 | 1.49 | 17,527,500 |
| 14 | 03/01/2021 | 28 | TAUFIQ HIDAYATILLAH | 280 | 400 | 19,000 | 1.43 | 7,600,000 |
| 15 | 06/01/2021 | 149 | MALHAN ADY TAMA | 301 | 385.5 | 20,000 | 1.28 | 7,710,000 |
| | | | | | | | | |
| 3817 | 26/06/2023 | 44 | TRISNA SOMARA | 380 | 800 | 22,500 | 2.11 | 18,000,000 |
| 3818 | 26/06/2023 | 192 | TRIYA | 779 | 1600 | 22,300 | 2.05 | 35,680,000 |
| 3819 | 27/06/2023 | 44 | TRISNA SOMARA | 662 | 1418 | 22,800 | 2.14 | 32,330,400 |
| 3820 | 27/06/2023 | 44 | TRISNA SOMARA | 694 | 1500 | 22,800 | 2.16 | 34,200,000 |
| 3821 | 27/06/2023 | 44 | TRISNA SOMARA | 235 | 506.5 | 22,800 | 2.16 | 11,548,200 |

Data Processing

Pada tahap ini, merupakan proses untuk mengubah data ke dalam format yang lebih efektif dan sesuai untuk digunakan dalam proses segmentasi menggunakan model RFM. Berikut adalah hasil dari tahap data preprocessing yang akan ditampilkan pada tabel di bawah ini :

Tabel 2 Data Processing

| No | NAMA | Recency | Frequency | Monetary |
|----|-------------------------|---------|-----------|----------------|
| 1 | MUHAMAD LUTFI | 290 | 51 | 1,181,768,000 |
| 2 | MALHAN ADY TAMA | 767 | 79 | 1,103,716,000 |
| 3 | SUJAJA | 864 | 6 | 88,302,000 |
| 4 | YASIN | 7 | 1002 | 18,264,656,950 |
| 5 | TAUFIQ HIDAYATILLAH | 699 | 18 | 190,920,500 |
| 6 | DIAN TAMARA | 45 | 200 | 4,955,958,600 |
| 7 | ONO YULIANTO | 12 | 375 | 6,743,352,600 |
| 8 | IBRAHAM BRAHMEDI | 12 | 502 | 9,210,912,950 |
| 9 | TRISNA SOMARA | 3 | 863 | 13,905,115,300 |
| 10 | PUPUT RIYADI | 834 | 11 | 141,380,000 |
| 11 | SANIAH | 622 | 22 | 333,410,000 |
| 12 | BUDI RUSKAMTO | 12 | 167 | 2,785,601,750 |
| 13 | AL1 SATORI | 8 | 101 | 1,884,001,250 |
| 14 | ONO DARSONO | 13 | 96 | 1,956,995,250 |
| 15 | TOYIB | 13 | 133 | 2,070,756,500 |
| 16 | SUTANA | 573 | 18 | 381,346,100 |
| 17 | WAWAN HERMAWAN | 539 | 4 | 87,700,000 |
| 18 | MALUL AZAM | 12 | 64 | 1,483,504,400 |
| 19 | HERI KISWANTO | 13 | 45 | 1,131,053,400 |
| 20 | BUBUN HERMAWAN | 488 | 2 | 22,472,100 |
| 21 | MUHTAR | 486 | 2 | 59,385,000 |
| 22 | AHMAD RIVAI | 290 | 1 | 11,135,000 |
| 23 | ISKANDAR WIDI | 64 | 36 | 841,364,850 |
| 24 | KUSWANDI | 64 | 8 | 116,000,000 |
| 25 | RIB'I HUDZAIFAH AL FATH | 41 | 3 | 38,625,000 |
| 26 | MUHAMAD SUGANDHI | 84 | 1 | 10,500,000 |
| 27 | ADE SUPRIATNA | 63 | 1 | 30,600,000 |
| 28 | DIAN KUSDIANA | 62 | 5 | 37,800,000 |
| 29 | MEA GUNAWAN | 62 | 1 | 21,440,100 |
| 30 | AJI RAHAYU | 24 | 1 | 36,480,500 |
| 31 | OPI NOVIANA | 7 | 1 | 27,500,000 |
| 32 | TRIYA | 4 | 2 | 72,240,850 |

Data Transformation

Pada tahap ini, dilakukan transformasi data agar siap digunakan dalam fase pemodelan. Sebelumnya, terdapat perbedaan rentang nilai yang signifikan antar atribut, yang dapat

mempengaruhi hasil analisis data. Oleh karena itu, dilakukan normalisasi data menggunakan metode *Min-Max Normalization*. Berikut adalah hasil dari tahap normalisasi data dengan metode *Min-Max Normalization*:

Tabel 3 *Data Transformation*

| Nama | <i>Recency</i> | <i>Frequency</i> | <i>Monetary</i> |
|-------------------------|----------------|------------------|-----------------|
| MUHAMAD LUTFI | 0.333 | 0.050 | 0.064 |
| MALHAN ADY TAMA | 0.887 | 0.078 | 0.060 |
| SUJAJA | 1.000 | 0.005 | 0.004 |
| YASIN | 0.005 | 1.000 | 1.000 |
| TAUFIQ HIDAYATILLAH | 0.808 | 0.017 | 0.010 |
| DIAN TAMARA | 0.049 | 0.199 | 0.271 |
| ONO YULIANTO | 0.010 | 0.374 | 0.369 |
| IBRAHAM BRAHMEDI | 0.010 | 0.500 | 0.504 |
| TRISNA SOMARA | 0.000 | 0.861 | 0.761 |
| PUPUT RIYADI | 0.965 | 0.010 | 0.007 |
| SANIAH | 0.719 | 0.021 | 0.018 |
| BUDI RUSKAMTO | 0.010 | 0.166 | 0.152 |
| AL1 SATORI | 0.006 | 0.100 | 0.103 |
| ONO DARSONO | 0.012 | 0.095 | 0.107 |
| TOYIB | 0.012 | 0.132 | 0.113 |
| SUTANA | 0.662 | 0.017 | 0.020 |
| WAWAN HERMAWAN | 0.623 | 0.003 | 0.004 |
| MALUL AZAM | 0.010 | 0.063 | 0.081 |
| HERI KISWANTO | 0.012 | 0.044 | 0.061 |
| BUBUN HERMAWAN | 0.563 | 0.001 | 0.001 |
| MUHTAR | 0.561 | 0.001 | 0.003 |
| AHMAD RIVAI | 0.333 | 0.000 | 0.000 |
| ISKANDAR WIDI | 0.071 | 0.035 | 0.046 |
| KUSWANDI | 0.071 | 0.007 | 0.006 |
| RIB'I HUDZAIFAH AL FATH | 0.044 | 0.002 | 0.002 |
| MUHAMAD SUGANDHI | 0.094 | 0.000 | 0.000 |
| ADE SUPRIATNA | 0.070 | 0.000 | 0.001 |
| DIAN KUSDIANA | 0.069 | 0.004 | 0.001 |
| MEA GUNAWAN | 0.069 | 0.000 | 0.001 |

| | | | |
|-------------|-------|-------|-------|
| AJI RAHAYU | 0.024 | 0.000 | 0.001 |
| OPI NOVIANA | 0.005 | 0.000 | 0.001 |
| TRIYA | 0.001 | 0.001 | 0.003 |

Perhitungan Manual

Pada tahap ini dilakukan pengujian penghitungan manual pada pemodelan *K-means Clustering*, data yang digunakan adalah data yang telah ditransformasi pada tahap sebelumnya. Perhitungan dilakukan menggunakan rumus *ecludian distance* sebagai berikut :

$$d(x, y) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$$

Tahap pertama dilakukan penentuan jumlah *cluster* dan titik awal *centroid*

Tabel 4 *Centroid* Awal

| <i>Centroid</i> | x | y | z |
|-----------------|-------|-------|-------|
| C1 | 0.010 | 0.166 | 0.152 |
| C2 | 0.006 | 0.100 | 0.103 |
| C3 | 0.012 | 0.095 | 0.107 |

Langkah selanjutnya adalah mengukur jarak antara pusat *cluster* menggunakan rumus *ecludian distance*.

1. Menghitung jarak data pertama dengan pusat *cluster*

$$d1(1,1) \sqrt{(0,333 - 0,010)^2 + (0,050 - 0,166)^2 + (0,064 - 0,152)^2} = 0,354$$

2. Menghitung jarak data pertama dengan pusat *cluster* kedua

$$d1(1,2) \sqrt{(0,333 - 0,006)^2 + (0,050 - 0,100)^2 + (0,064 - 0,103)^2} = 0,333$$

3. Menghitung jarak data pertama dengan pusat *cluster* ketiga

$$d1(1,3) \sqrt{(0,333 - 0,012)^2 + (0,050 - 0,095)^2 + (0,064 - 0,107)^2} = 0,327$$

Tabel 5 Perhitungan manual *iterasi* 1

| Data Ke | C1 | C2 | C3 | <i>Cluster</i> |
|---------|-------|-------|-------|----------------|
| 1 | 0.354 | 0.334 | 0.328 | C3 |
| 2 | 0.886 | 0.883 | 0.877 | C3 |
| 3 | 1.013 | 1.004 | 0.998 | C3 |

| | | | | |
|----|-------|-------|-------|----|
| 4 | 1.190 | 1.271 | 1.272 | C1 |
| 5 | 0.824 | 0.812 | 0.806 | C3 |
| 6 | 0.129 | 0.200 | 0.198 | C1 |
| 7 | 0.300 | 0.382 | 0.383 | C1 |
| 8 | 0.486 | 0.567 | 0.568 | C1 |
| 9 | 0.924 | 1.007 | 1.008 | C1 |
| 10 | 0.978 | 0.968 | 0.962 | C3 |
| 11 | 0.736 | 0.722 | 0.717 | C3 |
| 12 | 0.000 | 0.083 | 0.084 | C1 |
| 13 | 0.083 | 0.000 | 0.009 | C2 |
| 14 | 0.084 | 0.009 | 0.000 | C3 |
| 15 | 0.052 | 0.034 | 0.037 | C2 |
| 16 | 0.681 | 0.667 | 0.661 | C3 |
| 17 | 0.650 | 0.632 | 0.626 | C3 |
| 18 | 0.125 | 0.043 | 0.041 | C3 |
| 19 | 0.152 | 0.070 | 0.068 | C3 |
| 20 | 0.596 | 0.575 | 0.570 | C3 |
| 21 | 0.594 | 0.573 | 0.567 | C3 |
| 22 | 0.394 | 0.357 | 0.352 | C3 |
| 23 | 0.179 | 0.108 | 0.104 | C3 |
| 24 | 0.224 | 0.149 | 0.146 | C3 |
| 25 | 0.225 | 0.146 | 0.144 | C3 |
| 26 | 0.240 | 0.168 | 0.165 | C3 |
| 27 | 0.232 | 0.156 | 0.153 | C3 |
| 28 | 0.229 | 0.153 | 0.150 | C3 |
| 29 | 0.232 | 0.156 | 0.153 | C3 |
| 30 | 0.224 | 0.143 | 0.142 | C3 |
| 31 | 0.224 | 0.143 | 0.142 | C3 |
| 32 | 0.222 | 0.140 | 0.140 | C3 |

Hasil *cluster* didapatkan dari nilai terkecil setelah dilakukan perhitungan jarak terhadap *centroid* awal, kemudian tahap selanjutnya menentukan *centroid* baru dari perhitungan pada *iterasi* pertama yang akan digunakan *iterasi* selanjutnya hingga hasil *cluster* tidak berubah .

Tabel 66 *Centroid* baru Iterasi 4 (Akhir)

| <i>Centroid</i> | X | Y | Z |
|-----------------|-------|-------|-------|
| C1 | 0.006 | 0.684 | 0.659 |
| C2 | 0.068 | 0.047 | 0.053 |
| C3 | 0.754 | 0.017 | 0.014 |

1. Menghitung jarak data dengan *centroid* iterasi 4

$$d1(1,1) \sqrt{(0,333 - 0,006)^2 + (0,050 - 0,684)^2 + (0,064 - 0,659)^2} = 0,928$$

2. Menghitung jarak data pertama dengan *centroid* baru 4

$$d1(1,2) \sqrt{(0,333 - 0,068)^2 + (0,050 - 0,047)^2 + (0,064 - 0,053)^2} = 0,265$$

3. Menghitung jarak data pertama dengan *centroid* baru 4

$$d1(1,3) \sqrt{(0,333 - 0,754)^2 + (0,050 - 0,017)^2 + (0,064 - 0,014)^2} = 0,425$$

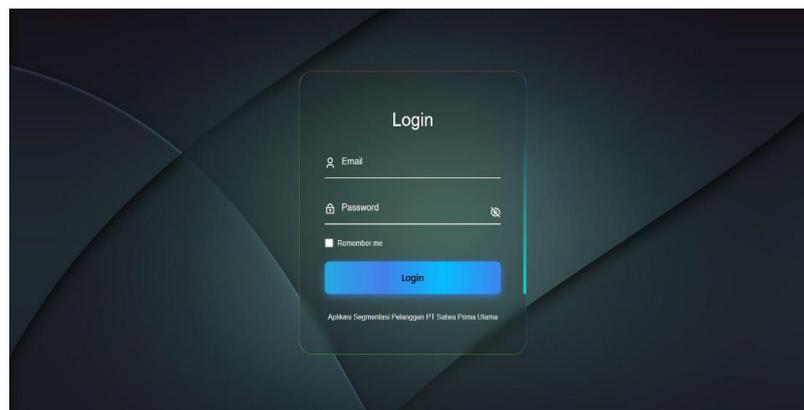
Tabel 7 Perhitungan manual iterasi 4 (Akhir)

| Data Ke | C1 | C2 | C3 | Cluster | Kategori |
|---------|-------|-------|-------|---------|------------------|
| 1 | 0.928 | 0.265 | 0.425 | C2 | Potensial |
| 2 | 1.225 | 0.82 | 0.153 | C3 | Tidak Potensial |
| 3 | 1.37 | 0.934 | 0.246 | C3 | Tidak Potensial |
| 4 | 0.465 | 1.345 | 1.581 | C1 | Sangat Potensial |
| 5 | 1.228 | 0.742 | 0.054 | C3 | Tidak Potensial |
| 6 | 0.622 | 0.266 | 0.773 | C2 | Potensial |
| 7 | 0.424 | 0.458 | 0.898 | C1 | Sangat Potensial |
| 8 | 0.24 | 0.642 | 1.013 | C1 | Sangat Potensial |
| 9 | 0.205 | 1.081 | 1.356 | C1 | Sangat Potensial |
| 10 | 1.341 | 0.899 | 0.211 | C3 | Tidak Potensial |
| 11 | 1.165 | 0.652 | 0.036 | C3 | Tidak Potensial |
| 12 | 0.724 | 0.165 | 0.771 | C2 | Potensial |
| 13 | 0.806 | 0.095 | 0.758 | C2 | Potensial |
| 14 | 0.807 | 0.091 | 0.752 | C2 | Potensial |
| 15 | 0.776 | 0.118 | 0.758 | C2 | Potensial |
| 16 | 1.132 | 0.596 | 0.092 | C3 | Tidak Potensial |
| 17 | 1.127 | 0.558 | 0.133 | C3 | Tidak Potensial |
| 18 | 0.848 | 0.066 | 0.748 | C2 | Potensial |
| 19 | 0.875 | 0.057 | 0.745 | C2 | Potensial |
| 20 | 1.1 | 0.5 | 0.192 | C3 | Tidak Potensial |
| 21 | 1.097 | 0.498 | 0.194 | C3 | Tidak Potensial |
| 22 | 1.004 | 0.275 | 0.422 | C2 | Potensial |
| 23 | 0.895 | 0.015 | 0.684 | C2 | Potensial |
| 24 | 0.942 | 0.062 | 0.684 | C2 | Potensial |
| 25 | 0.948 | 0.073 | 0.71 | C2 | Potensial |
| 26 | 0.953 | 0.076 | 0.661 | C2 | Potensial |
| 27 | 0.951 | 0.07 | 0.685 | C2 | Potensial |

| | | | | | |
|----|-------|-------|-------|----|-----------|
| 28 | 0.947 | 0.067 | 0.686 | C2 | Potensial |
| 29 | 0.951 | 0.071 | 0.686 | C2 | Potensial |
| 30 | 0.949 | 0.083 | 0.73 | C2 | Potensial |
| 31 | 0.949 | 0.095 | 0.75 | C2 | Potensial |
| 32 | 0.946 | 0.095 | 0.753 | C2 | Potensial |

Ini merupakan Iterasi terakhir dikarenakan pada iterasi ke 5 data hasil cluster sudah tidak berubah. Dari perhitungan manual ini didapatkan hasil terdiri dari 4 pelanggan sebagai kategori sangat potensial, 19 pelanggan sebagai kategori potensial, dan 9 pelanggan sebagai kategori tidak potensial.

Tampilan Program



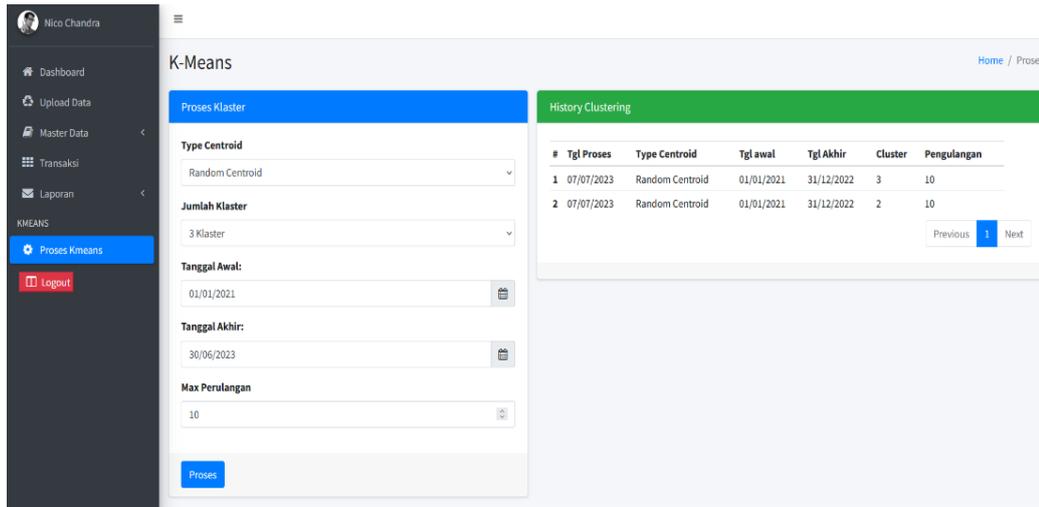
Gambar 2 Halaman *Login*

Halaman ini merupakan halaman awal yang menampilkan halaman login user yang diisi email dan sandi yang terdaftar.

| # | No Faktur | Tanggal | Kode Pelanggan | Nama Pelanggan | Ekor | Qty (Kg) | Harga | Bwa | Total | Aksi |
|----|-----------|------------|----------------|----------------|------|----------|--------|------|------------|----------------|
| 1 | 050491 | 27-06-2023 | 044 | TRISNA SOMARA | 662 | 1,418.00 | 22,800 | 0.00 | 32,330,400 | [Edit] [Hapus] |
| 2 | 050492 | 27-06-2023 | 044 | TRISNA SOMARA | 694 | 1,500.00 | 22,800 | 2.16 | 34,200,000 | [Edit] [Hapus] |
| 3 | 050493 | 27-06-2023 | 044 | TRISNA SOMARA | 235 | 506.50 | 22,800 | 2.16 | 11,548,200 | [Edit] [Hapus] |
| 4 | 050487 | 26-06-2023 | 044 | TRISNA SOMARA | 574 | 1,200.00 | 22,500 | 2.09 | 27,000,000 | [Edit] [Hapus] |
| 5 | 050488 | 26-06-2023 | 044 | TRISNA SOMARA | 483 | 1,000.00 | 22,500 | 2.07 | 22,500,000 | [Edit] [Hapus] |
| 6 | 050489 | 26-06-2023 | 044 | TRISNA SOMARA | 380 | 800.00 | 22,500 | 2.11 | 18,000,000 | [Edit] [Hapus] |
| 7 | 050490 | 26-06-2023 | 192 | TRIYA | 779 | 1,600.00 | 22,300 | 2.05 | 35,680,000 | [Edit] [Hapus] |
| 8 | 050482 | 25-06-2023 | 044 | TRISNA SOMARA | 287 | 550.00 | 22,500 | 1.92 | 12,375,000 | [Edit] [Hapus] |
| 9 | 050483 | 25-06-2023 | 192 | TRIYA | 846 | 1,639.50 | 22,300 | 1.94 | 36,560,850 | [Edit] [Hapus] |
| 10 | 050484 | 25-06-2023 | 044 | TRISNA SOMARA | 845 | 1,250.00 | 22,500 | 1.48 | 28,125,000 | [Edit] [Hapus] |

Gambar 3 Halaman data Transaksi

Halaman ini merupakan halaman transaksi yang akan menampilkan seluruh data transaksi yang ada, pada halaman ini dapat dilakukan proses pencarian data transaksi, serta dapat menambah, mengedit, dan juga menghapus data transaksi.



Gambar 4 Halaman Proses K-Means Clustering

Halaman ini merupakan tampilan halaman proses *clustering* pelanggan , dan juga histori clustering.

Pengujian Sistem

Pengujian sistem dilakukan untuk proses uji coba sistem yang dijalankan pada aplikasi segmentasi pelanggan menggunakan model RFM dan algoritma *k-means* dan sebagai evaluasi yang dihasilkan oleh sistem ini . Data yang diuji merupakan data transaksi perusahaan periode januari 2021- Juni 2023 yang telah dilakukan pemodelan menggunakan model RFM dan normalisasi data menggunakan metode *min-max*.

| # | Customer | Recency | Frequency | Monetary | Recency | Frequency | Monetary | Cluster | Label |
|----|---------------------|---------|-----------|--------------------|---------|-----------|----------|-----------|------------------|
| 1 | IBRAHAM BRAHMEDI | 61 | 502 | 9,210,912,950.000 | 0.010 | 0.500 | 0.504 | Cluster 1 | Sangat Potensial |
| 2 | ONO YULIANTO | 61 | 375 | 6,743,352,600.000 | 0.010 | 0.374 | 0.369 | Cluster 1 | Sangat Potensial |
| 3 | TRISNA SOMARA | 52 | 863 | 13,905,115,300.000 | 0.000 | 0.861 | 0.761 | Cluster 1 | Sangat Potensial |
| 4 | YASIN | 56 | 1002 | 18,264,656,950.000 | 0.005 | 1.000 | 1.000 | Cluster 1 | Sangat Potensial |
| 5 | BUBUN HERMAWAN | 537 | 2 | 22,472,100.000 | 0.563 | 0.001 | 0.001 | Cluster 2 | Tidak Potensial |
| 6 | MALHAN ADY TAMA | 816 | 79 | 1,103,716,000.000 | 0.887 | 0.078 | 0.060 | Cluster 2 | Tidak Potensial |
| 7 | MUHTAR | 535 | 2 | 59,385,000.000 | 0.561 | 0.001 | 0.003 | Cluster 2 | Tidak Potensial |
| 8 | PUPUT RIYADI | 883 | 11 | 141,380,000.000 | 0.965 | 0.010 | 0.007 | Cluster 2 | Tidak Potensial |
| 9 | SANIAH | 671 | 22 | 333,410,000.000 | 0.719 | 0.021 | 0.018 | Cluster 2 | Tidak Potensial |
| 10 | SUJAJA | 913 | 6 | 88,302,000.000 | 1.000 | 0.005 | 0.004 | Cluster 2 | Tidak Potensial |
| 11 | SUTANA | 622 | 18 | 381,346,100.000 | 0.662 | 0.017 | 0.020 | Cluster 2 | Tidak Potensial |
| 12 | TAUFIQ HIDAYATILLAH | 748 | 18 | 190,920,500.000 | 0.808 | 0.017 | 0.010 | Cluster 2 | Tidak Potensial |
| 13 | WAWAN HERMAWAN | 588 | 4 | 87,700,000.000 | 0.623 | 0.003 | 0.004 | Cluster 2 | Tidak Potensial |

Gambar 5 Hasil Pengujian segmentasi sistem

| | | | | | | | | | |
|----|------------------------|-----|-----|-------------------|-------|-------|-------|-----------|-----------|
| 14 | ADE SUPRIATNA | 112 | 1 | 30,600,000.000 | 0.070 | 0.000 | 0.001 | Cluster 3 | Potensial |
| 15 | AHMAD RIVAL | 339 | 1 | 11,135,000.000 | 0.333 | 0.000 | 0.000 | Cluster 3 | Potensial |
| 16 | AJI RAHAYU | 73 | 1 | 36,480,500.000 | 0.024 | 0.000 | 0.001 | Cluster 3 | Potensial |
| 17 | DIAN KUSDIANA | 111 | 5 | 37,800,000.000 | 0.069 | 0.004 | 0.001 | Cluster 3 | Potensial |
| 18 | HERI KISWANTO | 62 | 45 | 1,131,053,400.000 | 0.012 | 0.044 | 0.061 | Cluster 3 | Potensial |
| 19 | ISKANDAR WIDI | 113 | 36 | 841,364,850.000 | 0.071 | 0.035 | 0.046 | Cluster 3 | Potensial |
| 20 | KUSWANDI | 168 | 8 | 116,000,000.000 | 0.135 | 0.007 | 0.006 | Cluster 3 | Potensial |
| 21 | MALUL AZAM | 61 | 64 | 1,483,504,400.000 | 0.010 | 0.063 | 0.081 | Cluster 3 | Potensial |
| 22 | MEA GUNAWAN | 111 | 1 | 21,440,100.000 | 0.069 | 0.000 | 0.001 | Cluster 3 | Potensial |
| 23 | MUHAMAD LUTFI | 339 | 51 | 1,181,768,000.000 | 0.333 | 0.050 | 0.064 | Cluster 3 | Potensial |
| 24 | MUHAMAD SUGANDHI | 133 | 1 | 10,500,000.000 | 0.094 | 0.000 | 0.000 | Cluster 3 | Potensial |
| 25 | OPI NOVIANA | 56 | 1 | 27,500,000.000 | 0.005 | 0.000 | 0.001 | Cluster 3 | Potensial |
| 26 | RIBI HUDZAIFAH AL FATH | 90 | 3 | 38,625,000.000 | 0.044 | 0.002 | 0.002 | Cluster 3 | Potensial |
| 27 | TRIYA | 53 | 2 | 72,240,850.000 | 0.001 | 0.001 | 0.003 | Cluster 3 | Potensial |
| 28 | ALI SATORI | 57 | 101 | 1,884,001,250.000 | 0.006 | 0.100 | 0.103 | Cluster 3 | Potensial |
| 29 | BUDI RUSKAMTO | 61 | 167 | 2,785,601,750.000 | 0.010 | 0.166 | 0.152 | Cluster 3 | Potensial |
| 30 | ONO DARSONO | 62 | 96 | 1,956,995,250.000 | 0.012 | 0.095 | 0.107 | Cluster 3 | Potensial |
| 31 | TOYIB | 62 | 133 | 2,070,756,500.000 | 0.012 | 0.132 | 0.113 | Cluster 3 | Potensial |
| 32 | DIAN TAMARA | 94 | 200 | 4,955,958,600.000 | 0.049 | 0.199 | 0.271 | Cluster 3 | Potensial |

Gambar 6 Hasil Pengujian segmentasi sistem

Dari hasil pengujian sistem menggunakan dataset maka didapatkan hasil sebagai berikut : Sistem mampu mengelompokan data transaksi pelanggan dalam 3 klaster yang terdiri dari total 32 pelanggan, terdiri dari 9 pelanggan sebagai kategori pelanggan yang tidak potensial , dan 19 pelanggan sebagai kategori pelanggan potensial , dan 4 pelanggan sebagai kategori pelanggan sangat potensial.

Simpulan

Berdasarkan implementasi pembuatan aplikasi segmentasi pelanggan menggunakan k-means dan model RFM yang telah dibuat dan dijelaskan pada penelitian ini maka dapat ditarik kesimpulan sebagai berikut :

- Aplikasi segmentasi ini dapat bekerja baik dalam mengelompokan data transaksi dalam menentukan segmentasi pelanggan berdasarkan uji coba aplikasi yang dilakukan oleh PT. Satwa Prima Utama.
- Algoritma K-means dengan model RFM berhasil di implementasikan dalam kedalam aplikasi ini untuk membantu perusahaan mendapatkan informasi yang berguna dalam menghadapi persaingan bisnis yang kompetitif
- Dengan mengoptimalkan strategi pemasaran dan manajemen pelanggan, PT. Satwa Prima Utama dapat meningkatkan efisiensi operasional secara keseluruhan. Penerapan aplikasi ini juga membuka peluang untuk mengenali potensi pertumbuhan bisnis baru dan menjaga daya saing perusahaan di pasar yang semakin kompetitif dalam industri peternakan.

Saran

Adapun saran untuk penelitian selanjutnya sehingga menjadi lebih baik untuk kedepannya:

- Penelitian selanjutnya diharapkan melakukan uji coba lebih lanjut dengan menggunakan berbagai dataset yang berbeda, termasuk dataset dengan ukuran dan karakteristik yang beragam seperti penambahan atribut yang lain
- Penelitian selanjutnya dapat menambahkan dan mengeksplorasi metode lain sebagai perbandingan efektivitas dalam meningkatkan hasil segmentasi pelanggan
- Informasi yang dihasilkan oleh aplikasi ini sebaiknya digunakan sebagai landasan untuk mengembangkan strategi pemasaran yang lebih responsif.

Ucapan Terima Kasih

Terima Kasih kepada PT. Satwa Prima Utama telah memberikan saya tempat melakukan penelitian.

Referensi

- [1] S. A. Perdana, S. F. Florentin and A. Santoso, "Analisi Segmentasi Pelanggan Menggunakan K-means Clustering Studi Kasus Aplikasi Afagift," *Sebatik Vol. 26 No. 2*, 2022.
- [2] P. A. Wicaksana, I. B. A. Swamardika and R. S. Hartati, "Literature Review Analisis Perilaku Pelanggan Menggunakan RFM Model," *Majalah Ilmiah Teknologi Elektro, Vol. 21, No.1*, 2022.
- [3] Y. I. Kurniawan, "Perbandingan Algoritma Naive Bayes Dan C.45 Dalam Klasifikasi Data Mining," *Jurnal Teknologi Informasi dan Ilmu Komputer Vol.5, No.4*, 2018.
- [4] I. Werdingsih, B. Nuqoba and Muhammadun, *Data Mining Menggunakan Android, Weka, dan SPSS*, Jakarta: Airlangga University Press, 2020.
- [5] B. S. Pranata and D. P. Utomo, "Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor, Vol. 1, No.2," *Bulletin of Information Technology*, Vols. Vol.1, No.2, 2020.
- [6] A. Zauardi and H. Suprayitno, "Analisa Karakteristik Kecelakaan Lalu Lintas di Jalan Ahmad Yani Surabaya melalui Pendekatan Knowledge Discovery in Database," *Jurnal Manajemen Aset Infrastruktur & Fasilitas*, Vols. Vol.2, No.1, 2018.
- [7] J. Nasir, "Penerapan Data Mining Clustering Dalam Mengelompokkan Buku Dengan Metode K-Means," *Jurnal SIMETRIS*, Vols. Vol.11, No.2, 2020.
- [8] R. T. Vuldari, *Data Mining : Teori dan Aplikasi Rapidminer*, Jakarta: Penerbit Gaya Media, 2017.
- [9] D. A. Nasution, H. H. Khotimah and N. Chamidah, "Perbandingan Normalisasi Data Untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Journal of Computer Engineering System and Science Vol. 4, No. 1*, 2019.
- [10] I. S. A. E. P. Beta Estri Adiana, "Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering," *JUTEI*, vol. 2, p. 1, 2018.
- [11] S. Wahyuni and Fahrullah, "Segmentasi Pelanggan Berdasarkan Analisis Recency, Frequency, Monetary Menggunakan Algoritma K-Means," *Jurnal Simantec*, vol. 12, p. 1, 2023.
- [12] S. Wijaya and S. Thio, "Implementasi Membership Card dan Pengaruhnya Dalam Meningkatkan Loyalitas Pengunjung Restoran di Surabaya," *Eksekutif*, vol. 3, 2007.
- [13] F. Atik and A. P. Syahfar, "Segmentasi Konsumen Berdasarkan Model Recency, Frequency, Monetary dengan Metode K-Means," *JIEMS*, vol. 13 No.2, p. 1, 2020.
- [14] V. Virtusena, A. Johar and A. Wijanarko, "Pengelompokan Potensi Kelulusan Mahasiswa Fakultas Teknik Unib Menggunakan Algoritme K-Means," *Rekursif*, vol. 9, 2021.
- [15] E. Irfiani and S. S. Rani, "Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita," *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, vol. 6 no.4, p. 161, 2018.
- [16] Diotivani, I. Ruslianto And D. Prawira, "SISTEM SEGMENTASI PELANGGAN POTENSIAL MENGGUNAKAN METODE K-MEANS DAN ANALISIS RFM," *Coding : Jurnal Komputer Dan Aplikasi*, Vol. 11 No.3, 2023.