

Pengembangan Sistem CRM Berbasis Data mining: Segmentasi Nasabah Dengan Algoritma K-Means

M Wahyu Setiawan¹, Edy Rahman Syahputra²

^{1,2} Sitem Informasi, Universitas Harapan

Email: ¹wahyu211020@gmail.com, ²ydeaja@yahoo.com

Email Penulis Korespondensi: wahyu211020@gmail.com

Abstrak

Industri asuransi mobil di era transformasi digital menghadapi tantangan dalam mempertahankan loyalitas nasabah, terutama pada proses perpanjangan polis (renewal notice). Banyak nasabah yang menunda atau mengabaikan perpanjangan polis, sehingga berisiko kehilangan perlindungan asuransi dan berdampak pada menurunnya pendapatan perusahaan. Pendekatan Customer Relationship Management (CRM) konvensional belum mampu memberikan analisis mendalam terkait perilaku nasabah, sehingga perusahaan membutuhkan metode analitik berbasis data untuk mendukung strategi retensi yang lebih efektif. Penelitian ini mengusulkan pengembangan sistem Analytical CRM berbasis data mining dengan menggunakan algoritma K-Means Clustering pada PT Zurich Asuransi Indonesia. Algoritma ini digunakan untuk melakukan segmentasi nasabah berdasarkan variabel seperti rasio klaim, premi, usia kendaraan, serta riwayat perpanjangan polis. Segmentasi tersebut memberikan dasar bagi perusahaan dalam menyusun strategi komunikasi, promosi, dan layanan yang lebih personal dan relevan dengan kebutuhan nasabah asuransi mobil. Dengan demikian, sistem CRM berbasis data mining dapat mengoptimalkan proses renewal notice serta meningkatkan Customer Lifetime Value (CLV). Hasil penelitian menunjukkan bahwa penerapan algoritma K-Means mampu menghasilkan segmentasi nasabah yang lebih terstruktur dan objektif. Sistem CRM yang dibangun dapat mengolah data nasabah menjadi informasi strategis yang mendukung pengambilan keputusan perusahaan. Segmentasi yang dihasilkan juga membantu PT Zurich Asuransi Indonesia dalam merancang produk, strategi pemasaran, serta pelayanan yang lebih tepat sasaran untuk meningkatkan loyalitas nasabah asuransi mobil.

Kata Kunci: Data Mining, Algoritma Apriori, Peminatan Program Studi, Asosiasi Mata Kuliah, Optimalisasi Kurikulum

Abstract

Mistakes in choosing a study program specialization by students are a problem that can affect their academic results and academic future. This research aims to overcome this problem by applying data mining methods using the Apriori algorithm in the context of identifying association patterns between courses in the Informatics Engineering study program at Harapan University, Medan, Faculty of Engineering and Computers. By using this algorithm, this research provides a solution for designing a curriculum that is more efficient and relevant to the needs of students and industry. The research results show that the Apriori algorithm is able to identify patterns of association between course choice and student interest with a high level of confidence. An example is the correlation between student standards and the choice of certain study programs such as Multimedia and Networking. This analysis provides valuable insight for universities in optimizing curriculum preparation to better suit students' interests and needs, thereby improving the quality of higher education. By using the Apriori formula, the algorithm steps are carried out by looking for Itemset Frequency, Support, Confidence, Pruning, and Generate Association Rules. This research found an association between stambuk and students' choice of study program at Harapan University, Medan, Faculty of Engineering and Computers. For example, there is a significant association between the 2021 standard and the Multimedia study program, where 40% of students with the 2021 standard also chose the Multimedia study program, with a confidence level of 76.67%.

Keywords: Data Mining, Apriori Algorithm, Study Program Interest, Course Association, Curriculum Optimization

1. PENDAHULUAN

Di era transformasi digital, industri asuransi kendaraan bermotor menghadapi tantangan kompleks, khususnya dalam mempertahankan loyalitas nasabah melalui layanan yang berkelanjutan dan terpersonalisasi. Salah satu tantangan utama yang dihadapi perusahaan asuransi ialah bagaimana mengelola proses perpanjangan polis (renewal notice) secara efektif. Nasabah sering kali mengabaikan atau terlambat dalam melakukan perpanjangan polis, yang berpotensi menyebabkan kehilangan perlindungan asuransi dan berdampak pada pendapatan perusahaan [1].

Optimalisasi strategi manajemen hubungan pelanggan menjadi hal yang penting dalam meningkatkan efektivitas perpanjangan polis asuransi kendaraan, karena berpengaruh langsung terhadap tingkat retensi nasabah dan stabilitas pendapatan perusahaan. Pendekatan CRM tradisional belum mampu menyediakan analisis perilaku nasabah secara komprehensif, sehingga diperlukan penerapan metode analitik berbasis data seperti algoritma K-Means untuk menghasilkan segmentasi yang lebih akurat. Dengan segmentasi yang tepat, perusahaan dapat merancang strategi komunikasi dan pelayanan yang lebih personal, adaptif, dan berbasis data, sesuai dengan tuntutan transformasi digital dalam industri asuransi.

Customer Relationship Management (CRM) menjadi pendekatan strategis dalam menjalin hubungan jangka panjang dengan nasabah, termasuk dalam mengelola proses perpanjangan polis. Namun, sistem CRM konvensional cenderung bersifat administratif dan tidak mampu memberikan wawasan mendalam mengenai perilaku nasabah terkait perpanjangan polis, seperti frekuensi perpanjangan, alasan tidak memperpanjang, atau potensi loyalitas jangka panjang [2]. Untuk itu,

dibutuhkan sistem CRM yang bersifat analitis (Analytical CRM), yang mampu mengolah dan menganalisis data historis nasabah untuk mengidentifikasi pola-pola perilaku dalam proses perpanjangan polis. Dengan penerapan data mining sebagai alat analisis, Analytical CRM dapat memberikan informasi yang lebih tajam dan berbasis data, sehingga membantu perusahaan dalam menyusun strategi pengingat, promosi, dan personalisasi layanan dalam konteks renewal notice.

Salah satu metode yang relevan dalam pengembangan Analytical CRM ialah data mining, khususnya algoritma K-Means Clustering. Algoritma ini mampu mengelompokkan nasabah ke dalam segmen-segmen berdasarkan atribut tertentu, seperti frekuensi perpanjangan polis, rasio klaim, jumlah premi, dan usia kendaraan. Segmentasi ini dapat membedakan antara nasabah loyal, nasabah berisiko tinggi, hingga nasabah baru yang membutuhkan pendekatan khusus. Dengan demikian, perusahaan dapat merancang strategi komunikasi dan pelayanan yang lebih personal, adaptif, dan efisien [3]. PT Zurich Asuransi Indonesia kendaraan terkemuka di Indonesia, memiliki kepentingan strategis dalam meningkatkan rasio perpanjangan polis dan mengoptimalkan Customer Lifetime Value (CLV). Oleh karena itu, pengembangan sistem CRM berbasis data mining dengan pendekatan K-Means menjadi solusi yang relevan dalam meningkatkan efektivitas renewal notice serta memperkuat strategi retensi pelanggan.

Beberapa penelitian terdahulu menunjukkan bahwa integrasi antara sistem Customer Relationship Management (CRM) dan algoritma K-Means Clustering efektif dalam meningkatkan akurasi segmentasi pelanggan. Misalnya, penelitian oleh [4] menunjukkan peningkatan akurasi strategi promosi hingga 35% melalui segmentasi pelanggan ritel. Studi lain oleh [5] menerapkan variabel seperti rasio klaim, jumlah premi, dan jenis kendaraan dalam mengelompokkan nasabah asuransi, yang terbukti meningkatkan efisiensi layanan. Penelitian yang dilakukan oleh [3] menerapkan algoritma K-Means pada data nasabah asuransi di Indonesia dengan menggunakan variabel seperti umur kendaraan, umur nasabah, dan riwayat klaim atau pembayaran premi. Hasil segmentasi menunjukkan bahwa penggunaan variabel tersebut mampu memberikan wawasan strategis yang berguna dalam menyusun layanan dan pemasaran yang lebih tepat sasaran dan personal. Temuan-temuan ini menunjukkan relevansi penggunaan variabel rasio klaim, usia kendaraan, dan premi dalam proses segmentasi nasabah. Seluruh jurnal tersebut dapat diakses dan diunduh secara bebas melalui database jurnal ilmiah nasional atau repositori institusi pendidikan tinggi terkait.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Untuk mendukung proses pengembangan sistem CRM berbasis algoritma K-Means, data dikumpulkan melalui dua teknik utama, yaitu:

1. Wawancara (*Interview*)

Wawancara dilakukan kepada pihak internal PT Zurich Asuransi Indonesia yang berwenang dalam pengelolaan data nasabah dan strategi CRM. Tujuannya ialah untuk memperoleh pemahaman yang lebih dalam mengenai kebutuhan sistem CRM, proses segmentasi nasabah yang sedang berjalan, serta indikator-indikator penting yang digunakan dalam menilai karakteristik nasabah, seperti:

2. Riset

Riset dipakai sebagai instrumen utama untuk memperoleh data kuantitatif dari nasabah yang ialah nasabah PT Zurich Asuransi Indonesia. Riset ini disusun dalam bentuk pertanyaan tertutup dengan sub indikator sederhana, yaitu:

a. Rasio Klaim

Tahun pembelian kendaraan menunjukkan usia kendaraan saat ini.

Tabel 1 Indikator Rasio Klaim

Indikator	Kategori Skor	Keterangan	Nilai
Rasio Klaim	Sangat Baik	Jarang klaim (≤ 1 kali per tahun)	3
	Baik	Kadang klaim (2–3 kali per tahun)	2
	Cukup Baik	Sering klaim (> 3 kali per tahun)	1

b. Usia Kendaraan

Tahun pembelian kendaraan menunjukkan usia kendaraan saat ini.

Tabel 2 Indikator Usia Kendaraan

Indikator	Kategori Skor	Keterangan	Nilai
Usia Kendaraan	Sangat Baik	0 – 2 tahun	3
	Baik	3 – 5 tahun	2
	Cukup Baik	> 5 tahun	1

c. Jumlah Premi (Per Tahun / Per Bulan)

Besaran premi yang dibayarkan dan persepsi nasabah terhadap nilai manfaatnya.

Tabel 3 Indikator Jumlah Premi

Indikator	Kategori Skor	Keterangan	Nilai
Jumlah Premi	Sangat Baik	> Rp5.000.000	3
	Baik	Rp2.500.000 – Rp5.000.000	2
	Cukup Baik	< Rp2.500.000	1

Data yang diperoleh dari 96 nasabah kemudian direkapitulasi dan disajikan dalam bentuk tabel (Tabel 3.1) untuk dianalisis lebih lanjut. Hasil dari riset ini menjadi dasar untuk mengetahui segmentasi nasabah yang diberikan oleh PT Zurich Asuransi Indonesia.

Tabel 4 Data Nasabah

No	Nama Nasabah	Rasio klaim	Usia kendaraan	Jumlah premi
1	Alam Syahputra Leowandi	Baik	Sangat Baik	Sangat Baik
2	Alfad Tahara Ramadhan Silalahi	Sangat Baik	Baik	Sangat Baik
3	Angga Kurniawan	Sangat Baik	Baik	Sangat Baik
4	Aprianus Jebua	Sangat Baik	Baik	Sangat Baik
5	Aulya Thyberegina Br Tarigan	Sangat Baik	Baik	Sangat Baik
6	Daniel Agustinus Siregar	Sangat Baik	Baik	Sangat Baik
7	Dias Prasetyo Saragih	Sangat Baik	Baik	Sangat Baik
8	Dicha Syahbandi	Sangat Baik	Baik	Sangat Baik
9	Dimaz Pratama	Sangat Baik	Baik	Sangat Baik
10	Dwi Nova Kembara Gule	Sangat Baik	Baik	Sangat Baik
...	...	...	...	...
96	Sapta Aulia	Baik	Baik	Baik

3. Studi Kepustakaan (*Library Research*)

Studi kepustakaan ialah metode pengumpulan data yang dilakukan dengan menelaah berbagai sumber literatur yang relevan guna mendukung landasan teori dalam penelitian. Peneliti mengkaji referensi yang berasal dari jurnal internasional, jurnal nasional, jurnal lokal, serta buku-buku ilmiah untuk memperoleh pemahaman mendalam terkait konsep, teori, dan hasil penelitian sebelumnya yang berkaitan dengan segmentasi nasabah, layanan asuransi, dan sistem informasi. Studi ini menjadi dasar dalam merumuskan kerangka konseptual serta acuan dalam merancang sistem dan analisis yang dilakukan.

4. Observasi

Metode observasi dilakukan untuk mengumpulkan data secara langsung dengan mengamati aktivitas dan proses administrasi yang berkaitan dengan nasabah di PT Zurich Asuransi Indonesia. Observasi difokuskan pada proses pelayanan, pengelolaan data nasabah, pengajuan klaim, dan interaksi antara perusahaan dan nasabah. Tujuan dari observasi ini ialah untuk memperoleh gambaran nyata mengenai kondisi sistem yang berjalan, sehingga dapat diidentifikasi kebutuhan, permasalahan, dan potensi pengembangan sistem CRM berbasis segmentasi nasabah.

2.2 Data Mining

Data mining merupakan proses penggalian informasi dan pola yang bermanfaat dari suatu data yang sangat besar. Proses *data mining* terdiri dari pengumpulan data, ekstraksi data, analisa data, dan statistik data [6]. Ia juga umum dikenal sebagai *knowledge discovery*, *knowledge extraction*, *data/pattern analysis*, *information harvesting*, dan lainnya. Empat proses dalam *data mining* ini akan menghasilkan model/ pengetahuan yang sangat berguna [7].

Data mining juga dapat didefinisikan sebagai penguraian kompleks dari sekumpulan data menjadi informasi yang memiliki potensi secara implisit (tidak nyata/jelas) yang sebelumnya belum diketahui. Ia juga dapat didefinisikan sebagai penggalian dan analisis dengan menggunakan peralatan otomatis atau semi otomatis, dari sebagian besar data yang memiliki tujuan yaitu menemukan pola yang memiliki arti atau maksud. *Data mining* termasuk ke dalam *knowledge discovery* di dalam *database* (KDD) [8].

2.3 Algoritma Data

Algoritma data adalah serangkaian instruksi atau prosedur yang dirancang untuk mengolah data dan mendapatkan informasi yang berharga dari data tersebut. Berbagai algoritma data digunakan dalam analisis data untuk menemukan pola, hubungan, atau wawasan yang mungkin tersembunyi dalam data. Contoh algoritma data yang umum digunakan meliputi regresi linier, k-means clustering, Naive Bayes, Apriori, decision tree, dan sebagainya [9].

Algoritma Apriori bekerja dalam beberapa tahapan. Pertama, ia mengidentifikasi *item* tunggal yang memiliki *support* di atas batas tertentu yang telah ditentukan. Selanjutnya, algoritma ini menggabungkan *item-item* tersebut untuk

membentuk *Itemset* yang lebih besar, dan kemudian mengukur *support* dan *confidence* untuk setiap *Itemset* yang dihasilkan. *Itemset* yang memenuhi batas *support* dan *confidence* akan menjadi aturan asosiasi yang relevan dalam dataset tersebut. Algoritma Apriori sangat berguna dalam berbagai konteks, seperti analisis pembelian konsumen, rekomendasi produk, dan manajemen stok toko, karena dapat membantu dalam mengidentifikasi asosiasi yang mungkin tidak terlihat secara langsung dalam data mentah, membantu organisasi membuat keputusan yang lebih cerdas dan strategis [10].

2.4 K-Means

Algoritma K-means ialah salah satu algoritma dengan partitional, karena K-Means didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai centroid awalnya. Algoritma K-means menggunakan proses secara berulang-ulang untuk mendapatkan basis data cluster. Dibutuhkan jumlah cluster awal yang diinginkan sebagai masukan dan menghasilkan titik centroid akhir sebagai output.

Metode K-means akan memilih pola k sebagai titik awal centroid secara acak atau random. Jumlah iterasi untuk mencapai cluster centroid akan dipengaruhi oleh calon cluster centroid awal secara random. Sehingga didapat cara dalam pengembangan algoritma dengan menentukan centroid cluster yang dilihat dari kepadatan data awal yang tinggi agar mendapatkan kinerja yang lebih tinggi. Hung et. al. menggunakan algoritma K-Means ini dalam penelitian mereka untuk meningkatkan kecepatan metode clustering dengan cara membuat penentuan centroid yang lebih sederhana (Indraputra & Fitriana, 2020).

Langkah-langkah melakukan *clustering* dengan metode *K-means* ialah sebagai berikut (Dinata, Safwandi, Hasdyna, & Azizah, 2020):

1. Menentukan nilai k sebagai jumlah cluster yang ingin di bentuk.
2. Menentukan nilai acak atau random untuk pusat cluster awal centroid sebanyak k, untuk menghitung jarak setiap data input terhadap masing-masing centroid dengan menggunakan rumus jarak Euclidean Distance yaitu :

$$d(x_i, \mu_j) = \sqrt{\sum_{l=1}^n (x_{il} - \mu_{jl})^2}$$

Dimana:

x_i = data ke-i
 μ_j = centroid pada cluster ke-j
 n = jumlah fitur
 l = indeks fitur

3. Mengelompokkan setiap data berdasarkan kedekatannya dengan *centroid* atau mencari jarak terkecil.
4. Memperbaharui nilai centroid baru, nilai centroid baru di peroleh dari rata-rata cluster yang bersangkutan dengan menggunakan rumus yaitu :

$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in s_j} x_j$$

Keterangan :

$\mu_j(t+1)$ = centroid baru pada iterasi (t+1)
 N_{sj} = Data pada cluster S_j

5. Apabila data setiap cluster belum berhenti, lakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap cluster tidak ada yang berubah.

2.5 CRM

Customer Relationship Management (CRM) ialah suatu pendekatan manajemen yang memungkinkan perusahaan untuk mengidentifikasi, menarik, dan meningkatkan hubungan dengan pelanggan. Customer Relationship Management (CRM) ialah sebuah pendekatan baru dalam mengelola hubungan korporasi dan pelanggan pada level bisnis sehingga dapat memaksimalkan komunikasi, pemasaran melalui pengelolaan berbagai kontak yang berbeda dengan pelanggan. Pendekatan ini memungkinkan untuk mempertahankan pelanggan dan memberikan nilai tambah terus menerus pada pelanggan, selain juga memperoleh keuntungan yang berkelanjutan

Customer Relationship Management dirancang untuk mengoptimalkan keuntungan, pendapatan, dan kepuasan pelanggan. Yang dimana dalam mengembangkan perilaku memuaskan pelanggan dengan cara menghubungkan proses dari pemasok ke pelanggan. Pelanggan menciptakan persepsi tentang perusahaan berdasarkan pengalaman mereka pada saat berinteraksi, pelangganpun harus dilibatkan dalam proses Customer Relationship Management untuk memberikan nilai yang berkelanjutan.

Tujuan dari Customer Relationship Management yaitu menggunakan hubungan dengan pelanggan untuk meningkatkan keuntungan perusahaan, menggunakan informasi untuk memberikan pelayanan yang memuaskan,

mendukung proses penjualan berulang kepada pelanggan. Indikator Customer Relationship Management (CRM). Customer Relationship Management ialah suatu metode dalam menarik perhatian, pemeliharaan kepuasan nasabah serta meningkatkan dan mempererat hubungan dengan pelanggan. Lebih lanjut, Customer Relationship Marketing (CRM) menyediakan data dan informasi terkait dengan pelanggan, antara lain perilaku dalam berbelanja, kebiasaan dalam mengkonsumsi produk, dan lain-lain.

3. HASIL DAN PEMBAHASAN

3.1 Penerapan K-Means

Penerapan metode K-Means dilakukan sebagai upaya untuk menyelesaikan permasalahan dalam perancangan sistem data mining, khususnya dalam pengelompokan (klasterisasi) data segementasi nasabah di PT Zurich Asuransi Indonesia. Metode ini bertujuan untuk mengelompokkan nasabah ke dalam beberapa segmen berdasarkan variabel-variabel tertentu seperti jumlah premi, usia kendaraan, dan rasio klaim.

Data yang terdapat dalam tabel data 3. 1 di atas perlu disiapkan untuk diolah dengan menggunakan algoritma K-Means. Untuk itu, data tersebut harus dinormalisasi atau dihapus bagian yang tidak diperlukan serta diinisialisasi. Berikut ini ialah langkah-langkah dalam algoritma K-Means dengan menetapkan jumlah cluster, misalnya $k = 3$.

Tabel 5 Konversi Data Penilaian

No	Nama Nasabah	Rasio klaim	Usia kendaraan	Jumlah premi
1	Alam Syahputra Leowandi	2	3	3
2	Alfad Tahara Ramadhan Silalahi	3	2	3
3	Angga Kurniawan	3	2	3
4	Aprianus Jebua	3	2	3
5	Aulya Thyberegina Br Tarigan	3	2	3
6	Daniel Agustinus Siregar	3	2	3
7	Dias Prasetyo Saragih	3	2	3
8	Dicha Syahbandi	3	2	3
9	Dimaz Pratama	3	2	3
10	Dwi Nova Kembara Gule	3	2	3
...	...	...	...	...
96	Sapta Aulia	2	2	2

Tabel 5 menunjukkan hasil pengelompokan data penilaian orang tua memakai algoritma K-Means Clustering. Adapun tabel *cluster* uraian dibawah ini:

Tabel 6 Cluster

No	Cluster	Keterangan
1	K-1	Cukup Baik
2	K-2	Baik
3	K3	Sangat Baik

Dalam menentukan *centroid* setiap *cluster* yang diambil dari data sumber berdasarkan pilihan secara acak atau *random*

Tabel 7 Data *Centroid* Awal

No	Centroid	Rasio klaim	Usia kendaraan	Jumlah premi
1	Centroid 1	2	3	3
2	Centroid 2	2	2	3
3	Centroid 3	1	1	1

Hitung jarak data ke *Centroid* memakai rumus *Euclidean*, data tersebut dari *Cluster* terdekatnya.

1. Jarak antara Segementasi nasabah nomor pertama dengan titik m1

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\
 &= \sqrt{(2 - 2)^2 + (3 - 3)^2 + (3 - 3)^2} \\
 &= 0
 \end{aligned}$$

2. Jarak antara Segementasi nasabah nomor pertama dengan titik m2

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\
 &= \sqrt{(2 - 2)^2 + (3 - 2)^2 + (3 - 3)^2} \\
 &= 1,000
 \end{aligned}$$

3. Jarak antara Segementasi nasabah nomor pertama dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2-1)^2 + (3-1)^2 + (3-1)^2}$$

$$= 3,000$$

Untuk lebih lengkapnya jarak pada setiap baris data, hasilnya seperti pada tabel berikut

Tabel 8 Hasil Perhitungan Iterasi Ke 1

No	DM1	DM2	DM3	Jarak Terdekat	Nilai Terdekat
1	0,000	1,000	3,000	C1	0,000
2	1,414	1,000	3,000	C2	1,000
3	1,414	1,000	3,000	C2	1,000
4	1,414	1,000	3,000	C2	1,000
5	1,414	1,000	3,000	C2	1,000
6	1,414	1,000	3,000	C2	1,000
7	1,414	1,000	3,000	C2	1,000
8	1,414	1,000	3,000	C2	1,000
9	1,414	1,000	3,000	C2	1,000
10	1,414	1,000	3,000	C2	1,000
...	...	...	...	...	...
96	1,414	1,000	1,732	C2	1,000
WCV = Total Nilai Terdekat					100,719

Dari tabel 8 adapun proses pencarian nilai WCV (*Within-Cluster Variability*) berikut :

$$\text{WCV} = (0,000 + 1,000 + 1,000 + 1,000 + \dots + 1,000)$$

$$= 100,719$$

Dari proses *cluster* tabel 9 di dapat Jumlah Tingkat KeBaikan nasabah sebagai berikut :

Tabel 9 Hasil Cluster 1, 2, dan 3 Iterasi-1

No	Cluster	Jumlah Cluster
1	Cluster-1	36
2	Cluster-2	58
3	Cluster-3	2

Dari tabel 9 adapun daftar hasil *cluster* anggota c1-c3 sebagai berikut

- C1 = { 1, 11, 12, 13, 14, 15, 16, 17, 18, 19, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 69, 70, 71, 72 }
- C2 = { 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 95, 96 }
- C3 = { 68, 94 }

Dari tabel 3.6 adapun proses pencarian nilai BCV (*Between-Cluster Variability*) berikut :

1. Jarak antara Segementasi nasabah m1 dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2-2)^2 + (3-2)^2 + (3-3)^2}$$

$$= 1,000$$

2. Jarak antara Segementasi nasabah m1 dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2-1)^2 + (3-1)^2 + (3-1)^2}$$

$$= 2,449$$

3. Jarak antara Segementasi nasabah m2 dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2-1)^2 + (2-1)^2 + (3-1)^2}$$

$$= 3,000$$

$$\text{BCV} = 1,000 + 2,449 + 3,000$$

$$= 6,449$$

Menghitung kembali nilai rasio dengan membandingkan nilai BCV dan WCV.

$$BCV/WCV = 6,449 / 100,719$$

$$= 0,064$$

Selanjutnya, pada tahap kedua iterasi, kita perlu menghitung lagi titik-titik pusat yang baru dari kelompok yang telah terbentuk dengan cara mencari nilai rata-rata dalam kelompok tersebut.

Diketahui :

K-1 = Sangat Baik

K-2 = Baik

K-3 = Cukup Baik

1. Pada *cluster* 1 terdapat 1 data maka nilai rata-rata pada *cluster*-1 ialah

$$K1 = 2 + 3 + 3 + \dots + 2 / 36 = 2,14$$

$$K2 = 3 + 3 + 3 + \dots + 3 / 36 = 2,60$$

$$K3 = 3 + 3 + 2 + \dots + 3 / 36 = 1,50$$

2. Pada *cluster* 2 terdapat 6 data maka nilai rata-rata pada *cluster*-2 ialah

$$K1 = 3 + 3 + 3 + \dots + 2 / 58 = 3,00$$

$$K2 = 2 + 2 + 2 + \dots + 2 / 58 = 1,66$$

$$K3 = 3 + 3 + 3 + \dots + 2 / 58 = 2,50$$

3. Pada *cluster* 3 terdapat 22 data maka nilai rata-rata pada *cluster*-3 ialah

$$K1 = 1 + 2 / 2 = 1,50$$

$$K2 = 3 + 2 / 2 = 2,50$$

$$K3 = 1 + 1 / 2 = 1,00$$

Tabel 10 Titik Pusat *Cluster* pada Iterasi-2

Titik Pusat Awal	Rasio klaim	Usia kendaraan	Jumlah premi
<i>Cluster</i> 1	2,14	3,00	1,50
<i>Cluster</i> 2	2,60	1,66	2,50
<i>Cluster</i> 3	1,50	2,50	1,00

Apabila data di setiap kelompok masih bergerak, ulangi dari langkah 2 hingga 5, sampai tidak ada perubahan pada anggota setiap kelompok. Jika nilai Centroid berbeda dari nilai Centroid sebelumnya, algoritma akan melanjutkan ke langkah selanjutnya. Setelah nilai k dan posisi awal kelompok diketahui, langkah selanjutnya ialah mengukur jarak antara pusat kelompok menggunakan jarak Euclidian, yang akan menghasilkan matriks jarak yaitu C1, C2, dan C3 seperti dijelaskan di bawah ini.

Tabel 11 Hasil Perhitungan Iterasi Ke 2

No	DM1	DM2	DM3	Jarak Terdekat	Nilai Terkecil
1	0,466	1,500	2,121	Sangat Baik	0,466
2	1,392	0,594	2,550	Baik	0,594
3	1,392	0,594	2,550	Baik	0,594
4	1,392	0,594	2,550	Baik	0,594
5	1,392	0,594	2,550	Baik	0,594
6	1,392	0,594	2,550	Baik	0,594
7	1,392	0,594	2,550	Baik	0,594
8	1,392	0,594	2,550	Baik	0,594
9	1,392	0,594	2,550	Baik	0,594
10	1,392	0,594	2,550	Baik	0,594
...	...	...	...	...	...
96	1,152	1,004	1,225	Baik	1,004
Total WCV					84,247

Dari tabel 11 adapun proses pencarian nilai WCV (*Within-Cluster Variability*) berikut :

$$WCV = (0,466 + 0,594 + 0,594 + 0,594 + \dots + 0,594)$$

$$= 84,247$$

Dari proses *cluster* tabel 12 di dapat Jumlah Tingkat KeBaikan nasabah sebagai berikut :

Tabel 12 Hasil *Cluster* 1, 2, dan 3 Iterasi-2

No	Cluster	Jumlah Cluster
1	Cluster-1	30
2	Cluster-2	58

3	Cluster-3	8
---	-----------	---

Dari tabel 12 adapun daftar hasil *cluster* anggota c1-c3 sebagai berikut

1. C1 = { 1, 12, 13, 14, 15, 16, 17, 18, 19, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 69, 70, 71, 72 }
2. C2 = { 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 95, 96 }
3. C3 = { 62, 63, 64, 65, 66, 67, 68, 94 }

Dari tabel 3.7 adapun proses pencarian nilai BCV (*Between-Cluster Variability*) berikut :

1. Jarak antara Segementasi nasabah m1 dengan titik m2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2,14 - 2,60)^2 + (3,00 - 1,66)^2 + (2,56 - 2,72)^2}$$

$$= 1,433$$
2. Jarak antara Segementasi nasabah m1 dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2,14 - 1,50)^2 + (3,00 - 2,50)^2 + (2,56 - 1,00)^2}$$

$$= 1,754$$
3. Jarak antara Segementasi nasabah m2 dengan titik m3

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2,60 - 1,50)^2 + (1,66 - 2,50)^2 + (2,72 - 1,00)^2}$$

$$= 2,214$$

$$BCV = 1,433 + 1,754 + 2,214$$

$$= 5,402$$

Menghitung kembali nilai rasio dengan membandingkan nilai BCV dan WCV.

$$BCV/WCV = 5,402 / 84,247$$

$$= 0,064$$

Pengelompokan data nasabah dilakukan menggunakan algoritma K-Means berdasarkan variabel-variabel yang relevan dalam konteks asuransi, seperti usia kendaraan, tahun pembuatan, dan jumlah premi. Proses iterasi dilakukan hingga hasil konvergen, dan pengelompokan terbaik diperoleh pada iterasi ke-2 karena hasil pada iterasi ke-3 dan ke-4 tidak menunjukkan perubahan signifikan (stabil).

Tabel 13 Pengelompokan Hasil *Cluster* dari iterasi 2

Cluster	Nomor Nasabah
Cluster	1, 11-19, 46-61, 69-72
Sangat Baik	2-10, 20-45, 73-93, 95-96
Baik	62-68, 94

Ringkasan Segmentasi Nasabah:

Tabel 3.14 Hasil Segmentasi

Kategori	Jumlah Nasabah
Sangat Baik	30
Baik	58
Cukup Baik	8

Hasil pengelompokan menunjukkan mayoritas nasabah berada dalam kategori "Baik" (58 nasabah), diikuti oleh "Sangat Baik" (30 nasabah), dan "Cukup Baik" (8 nasabah). Ini mengindikasikan bahwa sebagian besar nasabah telah menunjukkan keBaikan yang stabil, sementara sebagian lainnya menunjukkan loyalitas tinggi dan potensi pengembangan. Temuan ini dapat dimanfaatkan PT Zurich Asuransi Indonesia untuk menyusun strategi retensi nasabah potensial, memperkuat program loyalitas, serta mengarahkan promosi dan edukasi kepada segmen dengan potensi peningkatan. Berikut ringkasan hasil penerapan CRM berdasarkan segmentasi:

Tabel 15 Hasil CRM berdasarkan Segmentasi Nasabah

Kategori Nasabah	Jumlah	Strategi CRM yang Diterapkan
Sangat Baik	30	Dimasukkan ke program loyalitas, reminder perpanjangan polis melalui email/SMS dengan pendekatan personal, serta penawaran reward khusus.

Baik	58	Diberikan pengingat perpanjangan polis bertahap (H-30, H-15, H-7), serta penawaran produk tambahan (cross-selling) untuk meningkatkan nilai premi.
Cukup Baik	8	Diprioritaskan untuk tindak lanjut intensif melalui telepon/kunjungan agen, diberikan edukasi manfaat polis, dan promo khusus untuk meningkatkan retensi.

Hasil pengelompokan nasabah dengan algoritma K-Means yang telah dilakukan pada iterasi ke-2 kemudian diintegrasikan ke dalam sistem Customer Relationship Management (CRM). Setiap nasabah diberi label kategori berdasarkan hasil segmentasi, yaitu Sangat Baik, Baik, dan Cukup Baik. Kategori ini selanjutnya digunakan oleh sistem CRM untuk menentukan strategi komunikasi, pelayanan, dan retensi yang berbeda pada masing-masing kelompok.

3.2 Implementasi Sistem

Tampilan antarmuka (interface) ialah bagian penting dalam aplikasi karena berfungsi sebagai media interaksi antara pengguna dengan sistem. Antarmuka ini harus bersifat user-friendly, mudah dipahami, serta responsif terhadap berbagai perangkat. Pada aplikasi ini, antarmuka dilengkapi dengan beberapa form utama, yaitu Form Login, Form Menu Utama, Form Data Nasabah, Form Data Centroid, Form Proses Algoritma k-means, serta Laporan Hasil Segmentasi.

1. Form Login

Form login digunakan untuk masuk kedalam sistem agar lebih aman dari *user-user* yang tidak bertanggung jawab sebelum masuk ke *form* utama. Pastikan form memiliki input untuk username dan password, serta tombol untuk mengirimkan data login. Anda dapat menambahkan atribut required pada input username dan password untuk memastikan kedua bidang tersebut tidak kosong sebelum mengirimkan formulir. Setelah pengguna mengisi formulir login dan mengirimkannya, data tersebut harus diproses untuk memverifikasi keaslian pengguna. Berikut ialah tampilan *form login* :



Gambar 1 Form Login

2. Form Menu Utama

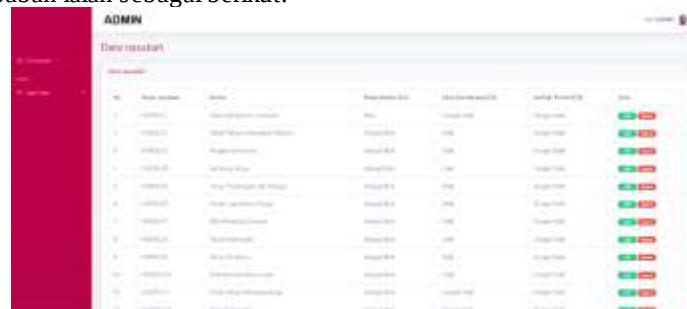
Form menu utama digunakan sebagai penghubung untuk *form* data nasabah, *form* data variabel, dan *form* proses algoritma k-means. Berikut ialah tampilan *form menu utama*:



Gambar 2 Form Menu Utama

1. Form Data Nasabah

Form data nasabah ialah form pengolahan data-data nasabah dalam penginputan data, ubah data dan penghapusan data. Adapun form nasabah ialah sebagai berikut.



Gambar 3 Form Data Nasabah

2. Form Data Centroid

Form data centroid ialah form pengolahan data-data centroid. Adapun form nasabah ialah sebagai berikut.



Gambar 4 Form Data Centroid

3. Form Proses Algoritma k-means

Form proses algoritma k-means ialah proses perhitungan dalam segmentasi data nasabah asuransi berdasarkan nasabah yang sudah ditentukan. Adapun Form proses algoritma k-means ialah sebagai berikut.

HASIL CLUSTERING		
Kode	Nama	Cluster
M00001	Alvin Spolitas Lasmadi	Sangat Baik
M00002	Andy Wahyu Setiawan	Sangat Baik
M00003	Chandi Samud	Sangat Baik
M00004	Edy Setiawan	Sangat Baik
M00005	Gilbert Nugroho	Sangat Baik
M00006	Herminia Ignatia Samud Sidiyasa	Sangat Baik
M00007	Herminia Ignatia Samud Sidiyasa	Sangat Baik
M00008	Irena Kurnia Ronggo	Sangat Baik
M00009	M. Rendi Sidiyasa	Sangat Baik
M00010	Miki Andy Kurniawan	Sangat Baik
M00011	Rizki Nurul Fathima	Sangat Baik
M00012	Uti Setiawan	Sangat Baik
M00013	Yusuf Setiawan Sidiyasa	Sangat Baik
M00014	Yusuf Setiawan Sidiyasa	Sangat Baik
M00015	Yusuf Setiawan Sidiyasa	Sangat Baik

Gambar 5 Form Proses Algoritma k-means



Gambar 6 Form Proses CRM

4. Laporan Hasil Segmentasi

Adapun hasil tersebut dalam bentuk laporan menampilkan hasil dalam bentuk laporan yang dapat dilihat pada Gambar 6 sebagai berikut:

9/1/25, 6:55 PM

PT Zurich Asuransi Indonesia

HASIL CLUSTERING

Kode	Nama	Cluster
NSB/01/1	Alam Syahputra Leonard	Sangat Baik
NSB/01/11	Feidy Yahya Simangunsong	Sangat Baik
NSB/01/12	Ghandi Gernadi	Sangat Baik
NSB/01/13	Gilang Triansyah	Sangat Baik
NSB/01/14	Gilbert Nanggolan	Sangat Baik
NSB/01/15	Immanuel Ignasius Samuel Siaban	Sangat Baik
NSB/01/16	Joanuel Agus Gabriel Lubis	Sangat Baik
NSB/01/17	Julio Kolode Hulapea	Sangat Baik
NSB/01/18	M. Farel Fadilah	Sangat Baik
NSB/01/19	Mhd. Andy Kurniawan	Sangat Baik
NSB/01/46	Egy Aditya Pratama	Sangat Baik
NSB/01/47	Ely Saltri	Sangat Baik
NSB/01/48	Faisal Arfanyah Lase	Sangat Baik

Gambar 7 Laporan Hasil Pengelompokan

4. KESIMPULAN

Berdasarkan hasil analisis dan implementasi, penerapan algoritma k-means pada sistem CRM terbukti mampu melakukan segmentasi nasabah secara lebih terstruktur dan objektif. Segmentasi ini membantu perusahaan dalam memahami karakteristik nasabah berdasarkan variabel yang dianalisis, sehingga proses identifikasi pola perilaku pelanggan dapat dilakukan dengan lebih akurat. Dengan demikian, sistem dapat menjadi alat yang efektif dalam mendukung pengelompokan nasabah sesuai dengan kebutuhan perusahaan.

Sistem CRM yang dibangun dengan integrasi algoritma k-means mampu mengolah data nasabah menjadi informasi yang bermanfaat. Informasi tersebut sangat penting bagi PT Zurich Asuransi Indonesia, khususnya dalam merumuskan strategi pemasaran dan pelayanan yang lebih tepat sasaran. Pemanfaatan hasil segmentasi ini memungkinkan perusahaan untuk meningkatkan efektivitas program promosi, serta memberikan layanan yang lebih sesuai dengan karakteristik tiap segmen pelanggan.

Namun, keberhasilan sistem ini sangat bergantung pada kualitas data yang digunakan. Data nasabah harus akurat dan lengkap, serta perlu dilakukan pengujian berulang agar hasil segmentasi benar-benar mencerminkan kondisi nyata. Dari hasil pengujian yang dilakukan, terlihat bahwa terdapat kelompok nasabah dengan kecenderungan tertentu yang dapat dijadikan dasar dalam perancangan produk asuransi, strategi promosi, dan peningkatan layanan. Hal ini menegaskan bahwa integrasi k-means dalam CRM mampu meningkatkan efektivitas pengambilan keputusan perusahaan.

UCAPAN TERIMA KASIH

Terima Kasih diucapkan kepada kedua orang tua serta keluarga yang selalu memberi motivasi, Doa dan dukungan moral maupun materi, serta pihak-pihak yang telah mendukung dalam proses pembuatan jurnal ini yang tidak dapat disebutkan satu persatu. Kiranya jurnal ini bisa memberi manfaat bagi pembaca dan dapat meningkatkan kualitas jurnal selanjutnya.

DAFTAR PUSTAKA

- [1] E. F. L. Awalina and W. I. Rahayu, "Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail," *Jurnal Teknologi dan Informasi (JATI)*, pp. 122-137, 2023.
- [2] S. F. Widyana and H. Firmansyah, "Pengaruh Customer Relationship Management (CRM) Terhadap Kepuasan Pelanggan Produk Sepatu Converse," *Jurnal Bisnis dan Pemasaran*, pp. 1-11, 2021.
- [3] A. A. A. Putri and S. A. Rahmah, "Implementasi Data Mining Dengan Algoritma K-Means Clustering Untuk Analisis Bisnis Pada Perusahaan Asuransi," *Djtechno : Jurnal Teknologi Informasi*, pp. 139-152, 2024.

- [4] V. Alvianatinova, I. Ali, N. Rahaningsih and A. Bahtiar, "Penerapan Algoritma K-Means Clustering Dalam Pengelompokan Data Penjualan Supermarket Berdasarkan Cabang (BRANCH)," *JATI (Jurnal Mahasiswa Teknik Informatika)*, pp. 1529-1535, 2024.
- [5] F. Razi, V. Terisia and D. Yusuf, "Penerapan Data Mining Untuk Menentukan Jenis Asuransi Yang Paling Diminati DI PT AIA Financial Dengan Metode Clustering," *Jurnal Ilmiah Teknologi dan Informasi*, pp. 1-13, 2023.
- [6] W. S. Ramadhan and N. Mandiri, "Implementasi Algoritma Apriori dalam Menentukan Pola Transaksi Penjualan," *Jurnal Infortech*, vol. VI, no. 1, pp. 52-58, 2024.
- [7] A. W. S, I. G. I. Sudipa, T. A. E. Putra, A. J. Wahidin, W. A. Syukrilla, A. K. Wardhani, N. Heryana, T. Indriyani and L. W. Santoso, *DATA MINING, Sumatera Barat: PT GLOBAL EKSEKUTIF TEKNOLOGI*, 2023.
- [8] M. Y. Ardilla, A. Manuhutu, N. Ahmad, I. Hasbi, G. M. A. Manuhutu, M. Ridwan, H. A. K. Wardhani, S. Alim, I. Romli, Y. Religia, D. T. Octafian, U. U. Sufandi and I. Ernawati, *Data Mining dan Aplikasinya*, Bandung: Widina Bhakti Persada, 2021.
- [9] P. M. S. Tarigan, J. T. Hardinata, H. Qurniawan, M. Safii and R. Winanjaya, "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus : Toko Sinar Harahap)," *UMJ*, p. 51 – 61, 2022.
- [10] S. and S. , "Penerapan Data Mining Denganmetode Algoritma Apriori Untukmenentukan Pola Pembelian Ikan," *Jurnal Sistem Informasi*, p. VI, 2019.
- [11] N. T. Ayu, J. Jasmir2 and I. S. Wijaya, "Penerapan Data Mining Menggunakan Algoritma Apriori Untuk Persediaan Stok Obat Pada Apotek Safa," *Jurnal Manajemen Teknologi dan Sistem Informasi (JMS)*, pp. 700-711, 2024.
- [12] A. Hanapi, R. Sari and M. , "Penerapan Algoritma Apriori Untuk Rekomendasi Produk Bagi Pelanggan Toko Online Berbasis Website," *Jurnal Jaring SainTek*, pp. 51-60, 2023.
- [13] H. Kusumo, E. Sedyono and M. Marwata, "Analisis Algoritma Apriori Untuk Mendukung Strategi Promosi Perguruan Tinggi," *Walisongo Journal of Information Technology*, pp. 51-62, 2019.
- [14] A. Zohriah, A. Muin and M. , "Paradigma Pendidikan di Era Digital," *JIIP (Jurnal Ilmiah Ilmu Pendidikan)* , pp. 4546-4554, 2023.
- [15] I. F. Jannah, "Penerapan Algoritma Apriori Pada Strategi Penjualan Di Giant," *KARISMATIKA*, pp. 1-10, 2021.
- [16] I. P. Astuti, "Algoritma Apriori Untuk Menemukan Hubungan Antara Jurusan Sekolah Dengan Tingkat Kelulusan Mahasiswa," *JURNAL TEKNIK INFORMATIKA*, pp. 69-78, 2019.
- [17] E. H. Pujiarini, "Analisis Asosiasi Untuk Menentukan Strategi Promosi Perguruan Tinggi Dengan Algoritma Apriori," *Jurnal Informatika dan Komputer (JIKO)* , pp. 45-51, 2019.