

# Clustering Loyalitas Pelanggan Menggunakan Algoritma K-Means Berbasis Web

**Diva Shafarina Aprilia, Sudriyanto, Moh Jasri\***

Program Studi Informatika, Fakultas Teknik, Universitas Nurul Jadid Paiton, Probolinggo

Email: <sup>1</sup>dievashafarina99@gmail.com, <sup>2</sup>sudriyanto@unuja.ac.id, <sup>3,\*</sup>jasri@unuja.ac.id

Email Penulis Korespondensi: jasri@unuja.ac.id

**Abstrak**-Kehadiran aktif pelanggan dalam bertransaksi dengan perusahaan berpengaruh besar pada profitabilitas. Pengelompokan data pelanggan menjadi loyal dan tidak loyal adalah metode umum untuk mengenali pola loyalitas, yang membantu perusahaan merancang strategi lanjutan seperti pemberian insentif sesuai tingkat kesetiaan pelanggan. Penerapan algoritma K-Means Clustering berbasis web memungkinkan manajemen PT Bhara Utama untuk mengakses hasil pengelompokan pelanggan dengan mudah, mempercepat analisis data, meningkatkan efisiensi pengambilan keputusan, dan memfasilitasi integrasi dengan sistem informasi yang ada. Eksperimen yang dilakukan pada 660 data pelanggan menghasilkan tiga kelompok: 8 pelanggan sangat loyal, 461 pelanggan cukup loyal, dan 191 pelanggan tidak loyal. Evaluasi akurasi menggunakan Davies-Bouldin Index (DBI) menunjukkan nilai sebesar 0,19, yang menunjukkan kualitas kluster yang baik.

**Kata Kunci:** Loyalitas Pelanggan; K-Means Clustering; Pengambilan Keputusan

**Abstract**-Active customer engagement in transactions with the company significantly impacts profitability. Categorizing customer data into loyal and non-loyal segments is a common method to identify loyalty patterns. The results of this segmentation can guide companies in designing follow-up strategies, including tailored incentives based on customer loyalty levels. Implementing a web-based K-Means Clustering algorithm allows PT Bhara Utama's management to easily access customer segmentation results, speeding up data analysis and enhancing decision-making efficiency. The use of web technology also facilitates integration with existing information systems and provides more flexible access. An experiment conducted on 660 customer data resulted in three groups: 8 very loyal customers, 461 moderately loyal customers, and 191 non-loyal customers. Accuracy evaluation using the Davies-Bouldin Index (DBI) showed a value of 0.19, indicating high-quality clusters.

**Keywords:** Customer Loyalty; K-Means Clustering; Decision-Making Efficiency

## 1. PENDAHULUAN

Teknologi semakin lama ada kemajuan, hal itu memungkinkan kita cepat dan tepat melakukan aktifitas kehidupan sehari-hari. Seiring berjalannya waktu, teknologi maju dengan sangat pesat. banyak Perusahaan dipaksa terus meningkatkan kemampuannya untuk mengolah informasi yang akurat untuk memenuhi keinginan, kebutuhan, serta kepuasan pelanggan agar tetap setia menggunakan produk Perusahaan masing-masing [1].

Dalam dunia perbisnisan sangat banyak persaingan dan pemasaran menjadi sangat ketat, oleh karena itu Perusahaan dituntut untuk meningkatkan kemampuannya dengan potensi yang ada. Pelanggan adalah salah satu potensi yang sangat penting dimiliki Perusahaan, oleh sebab itu saat ini Perusahaan lebih banyak memfokuskan aktifitas pemasaran pada mempertahankan dan meningkatkan loyalitas pelanggan lama dibandingkan untuk mencari pelanggan baru. Bukan berarti Perusahaan tidak membutuhkan pelanggan baru, hanya saja lebih ke memfokuskan kepada pelanggan lama karena biaya untuk mendapatkan pelanggan baru lebih banyak dibandingkan mempertahankan biaya pelanggan [2].

PT. Bhara Utama merupakan salah satu jenis usaha perdagangan yang bergerak dalam bidang E-Commerce yang menyediakan berbagai macam jenis pakaian. PT. Bhara Utama adalah sebuah toko yang menerima pembelian online, sebagian besar pelanggan berasal dari sumber online. Permasalahan toko Bhara Utama adalah seringkali tidak mengetahui pola penjualan yang terjadi karena toko Bhara Utama tidak memahami apa yang dibeli pelanggan. Dari data yang diperoleh dari penjualan toko Bhara Utama dapat digunakan untuk pengelompokan data pelanggan menjadi kategori loyal dan tidak loyal. agar penjualan dapat meningkat dengan menggunakan algoritma K-Means [3]. Eksplorasi ini merangkul metode dari algoritma K-Means, yang terkenal sebagai salah satu alat ampuh untuk mengurai masalah clustering dalam penambangan data. Algoritma ini menjalankan proses pengelompokan dengan menghitung nilai-nilai dan mengelompokkan objek-objek yang memiliki kedekatan satu sama lain.

Metode pengelompokan data K-Means Cluster ini bekerja dengan cara yang ajaib: pertama-tama, ia memilih titik tengah (centroid) untuk setiap kelompok. Data yang berkeliaran di sekitar dan merasa mirip dengan titik ajaib ini akan segera bergabung dalam kelompoknya. Setiap data harus memilih timnya dengan hati-hati. Setelah semua centroid punya pengikut setia, kelompok-kelompok ini akan diacak kembali untuk menciptakan tim baru. Proses tarian ini berlanjut sampai para centroid merasa nyaman dan tetap di tempat mereka. Tahapan algoritma K-Means adalah sebagai berikut

Menetapkan kuantitas kelompok, menaruh data secara acak dalam kelompok, menghitung titik pusat (centroid) dengan mengambil rata-rata dari data dalam setiap kelompok, menghitung karakteristik masing-masing data dan mencocokkannya dengan titik pusat awal, menugaskan setiap data ke kelompok yang memiliki jarak terdekat, mengulang langkah ketiga hingga semua data menetap dalam kelompoknya dan titik pusat tidak bergerak lagi.

K-Means adalah salah satu pendekatan nonhierarki yang aneh untuk membagi data menjadi dua atau lebih kelompok yang unik. Teknik ini membagi data menjadi satu kelompok di mana semua datanya mirip, sedangkan data dengan ciri-ciri yang berbeda akan digabungkan ke dalam kelompok lain. Inti dari pengelompokan ini adalah untuk mengurangi nilai fungsi objektif yang telah ditentukan dalam proses tersebut. Biasanya, tujuan utamanya adalah untuk

menekan variasi dalam kelompok dan meningkatkan perbedaan antara kelompok-kelompok yang berbeda [4]. Algoritma K-Means Clustering adalah salah satu metode data mining yang efektif untuk mengelompokkan data berdasarkan kesamaan tertentu. Dengan algoritma ini, kita bisa membagi pelanggan PT. Bhara Utama Maron menjadi beberapa segmen berdasarkan tingkat loyalitas mereka. Pengelompokan ini memberikan wawasan mendalam tentang karakteristik setiap kelompok pelanggan, yang bisa digunakan untuk merancang strategi pemasaran yang lebih tepat sasaran.

Penerapan algoritma K-Means Clustering berbasis web memungkinkan manajemen PT. Bhara Utama mengakses hasil pengelompokan pelanggan dengan mudah. Hal ini tidak hanya mempercepat analisis data tetapi juga meningkatkan efisiensi dalam pengambilan keputusan. Penggunaan teknologi berbasis web juga memudahkan integrasi dengan sistem informasi yang sudah ada dan memungkinkan akses yang lebih fleksibel.

Evaluasi kinerja sebuah perusahaan sering kali didasarkan pada profitabilitas yang dihasilkan. Kehadiran pelanggan yang aktif melakukan transaksi dengan perusahaan dapat memberikan dampak besar terhadap profitabilitas tersebut. Mengelompokkan data pelanggan ke dalam kategori loyal dan tidak loyal merupakan metode umum untuk mengidentifikasi pola loyalitas. Hasil pengelompokan ini kemudian dapat menjadi panduan bagi perusahaan dalam merancang strategi lanjutan, termasuk pemberian insentif yang disesuaikan dengan tingkat loyalitas pelanggan. Dengan mengenali pola loyalitas pelanggan sejak dini, perusahaan dapat mengambil langkah-langkah yang tepat untuk mempertahankan dan meningkatkan hubungan dengan pelanggan yang loyal [5].

Loyalitas pelanggan merupakan salah satu aset terpenting dalam bisnis yang dapat menentukan keberlangsungan dan pertumbuhan perusahaan. Pengelolaan loyalitas pelanggan yang efektif memerlukan pemahaman mendalam tentang pola perilaku konsumen dan segmentasi pasar. Dalam konteks ini, algoritma clustering seperti K-Means memainkan peran krusial dalam membantu perusahaan untuk memahami dan mengelompokkan pelanggan berdasarkan kesamaan dalam perilaku dan preferensi mereka [5]. Algoritma K-Means adalah metode yang banyak digunakan dalam clustering karena kesederhanaannya dan kemampuannya dalam mengelompokkan data menjadi sejumlah cluster berdasarkan jarak centroid. Metode ini sangat berguna dalam analisis data pelanggan karena dapat mengidentifikasi pola yang mungkin tidak terlihat secara langsung dalam data [6]. Di PT. Bhara Utama Maron, analisis loyalitas pelanggan merupakan kunci untuk meningkatkan strategi pemasaran dan pelayanan pelanggan. Implementasi sistem berbasis web memungkinkan perusahaan untuk memproses data pelanggan secara real-time dan memanfaatkan hasil clustering untuk keputusan bisnis yang lebih baik. Penggunaan algoritma K-Means dalam sistem berbasis web memberikan keuntungan tambahan dalam hal skalabilitas dan efisiensi, karena proses clustering dapat dilakukan dengan cepat bahkan pada dataset yang besar [7]. Selain itu, integrasi dengan platform web mempermudah visualisasi hasil clustering sehingga dapat diakses oleh berbagai pemangku kepentingan dalam perusahaan.

Sejumlah penelitian telah menunjukkan bahwa algoritma K-Means efektif dalam segmentasi pelanggan dan analisis perilaku. Dalam studi yang dilakukan oleh [8], K-Means berhasil mengidentifikasi kelompok pelanggan dengan tingkat loyalitas yang berbeda, yang selanjutnya digunakan untuk merancang strategi pemasaran yang lebih terarah. Hasil dari clustering ini memberikan wawasan yang lebih baik tentang bagaimana pelanggan berinteraksi dengan produk atau layanan dan bagaimana perusahaan dapat meningkatkan kepuasan pelanggan melalui penyesuaian layanan. K-Means juga memungkinkan penentuan jumlah cluster yang optimal menggunakan metode seperti Elbow atau Silhouette Score [9], yang membantu dalam menghindari overfitting dan memastikan bahwa segmentasi yang dihasilkan akurat.

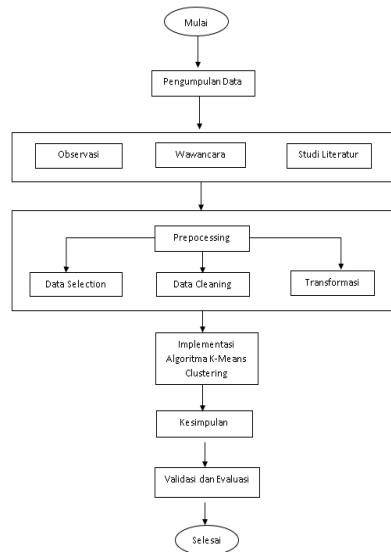
Keberhasilan implementasi K-Means dalam pengelolaan loyalitas pelanggan bergantung pada kualitas data yang digunakan dan parameter yang diatur dalam algoritma. Data yang bersih dan relevan sangat penting untuk mendapatkan hasil clustering yang akurat [10]. Penelitian lain menunjukkan bahwa pemilihan fitur yang tepat dan normalisasi data dapat mempengaruhi efektivitas K-Means dalam segmentasi pelanggan [11]. Selain itu, evaluasi hasil clustering harus dilakukan secara terus-menerus untuk memastikan bahwa model tetap relevan dengan perubahan dalam perilaku pelanggan dan kondisi pasar [12].

Sistem berbasis web memberikan fleksibilitas tambahan dalam penerapan algoritma K-Means karena dapat diakses dari berbagai lokasi dan perangkat, memungkinkan tim analisis untuk bekerja secara kolaboratif dan melakukan pembaruan secara berkala [13]. Teknologi ini juga mendukung integrasi dengan berbagai sumber data dan alat analisis lainnya [14]. Dengan menggunakan K-Means dalam lingkungan berbasis web, PT. Bhara Utama Maron dapat memanfaatkan potensi penuh dari data pelanggan untuk meningkatkan strategi bisnis dan membangun hubungan jangka panjang dengan pelanggan mereka [15]. Namun, tantangan yang dihadapi dalam penerapan algoritma K-Means termasuk penentuan jumlah cluster yang optimal dan penanganan data yang tidak seimbang [16]. Penelitian terbaru menunjukkan bahwa teknik seperti algoritma K-Means berbasis algoritma genetika atau metode hybrid dapat membantu mengatasi beberapa keterbatasan ini. Selain itu, adaptasi algoritma K-Means dengan teknik machine learning lainnya dapat meningkatkan hasil clustering dan analisis [17].

## **2. METODOLOGI PENELITIAN**

### **2.1 Kerangka Penelitian**

Rancangan penelitian adalah strategi atau proses yang direncanakan untuk mengumpulkan semua data sesuai dengan tujuan penelitian. Peneliti akan menggunakan metode kuantitatif dan model Knowledge Discovery in Database (KDD) dalam penelitian ini.



**Gambar 1.** Alur KDD

Penelitian ini menggunakan kerangka Knowledge Discovery in Database (KDD) sebagai landasan. Data mining, sebagai inti dari KDD, menerapkan algoritma untuk mengeksplorasi, membentuk model, dan menemukan pola tersembunyi dalam data yang disimpan secara elektronik. Proses KDD mencakup teknik analisis data canggih dan menarik perhatian dari berbagai bidang seperti kecerdasan buatan, machine learning, dan basis data. Algoritma self-learning dalam KDD berfungsi untuk menghasilkan informasi berharga dari data, melalui siklus berkelanjutan antara permintaan algoritma dan interpretasi pola yang saling memberikan umpan balik.

Penelitian ini mengklasifikasikan sumber data menjadi dua kategori: data primer dan data sekunder. Data primer diperoleh langsung oleh peneliti melalui berbagai metode seperti observasi, wawancara, dan survei, sehingga bersifat asli dan eksklusif. Dalam penelitian ini, data primer dikumpulkan melalui wawancara dengan pemilik PT. Bhara Utama Maron. Data sekunder, sebaliknya, berasal dari sumber yang lebih mudah diakses namun telah melalui beberapa tahap pengolahan statistik sehingga kehilangan sifat aslinya. Contoh data sekunder yang digunakan adalah data penjualan dari admin PT. Bhara Utama Maron yang mencakup atribut seperti ID customer, jumlah pesanan, harga satuan barang [18], total harga keseluruhan, dan frekuensi pembelian dalam setahun. Metode pengumpulan data dalam penelitian ini meliputi observasi langsung, wawancara, dan studi pustaka untuk mendapatkan informasi yang relevan dan mendalam.

Proses pengolahan data dalam penelitian ini melibatkan beberapa tahap, mulai dari preprocessing hingga analisis menggunakan algoritma K-Means Clustering. Preprocessing mencakup pengumpulan data mentah dan mengubahnya menjadi informasi yang berguna, serta pemilihan dan transformasi data ke dalam format yang lebih cocok untuk analisis data mining [19]. Algoritma K-Means digunakan untuk mengelompokkan data ke dalam sejumlah cluster berdasarkan jarak terdekat ke centroid, menggunakan metode Euclidean distance. Proses ini melibatkan iterasi berulang hingga pusat cluster tidak berubah lagi. Validasi dan evaluasi hasil clustering dilakukan untuk memastikan bahwa jumlah cluster yang diperoleh adalah yang terbaik, sehingga dapat memberikan informasi yang berharga bagi pengambilan keputusan [20].

### 3. HASIL DAN PEMBAHASAN

Hasil penelitian ini menggunakan dataset yang diperoleh dari laporan pembelian Pakaian di PT. Bhara Utama Maron. Parameter yang digunakan meliputi customer ID, jumlah pesanan, harga satuan, total keseluruhan, dan frekuensi pembelian. Algoritma K-Means Clustering bertujuan untuk mengelompokkan data dengan karakteristik yang mirip atau hampir sama ke dalam beberapa cluster. Dataset yang digunakan dalam percobaan penelitian ini adalah data pembelian dari tahun 2023-2024.

**Tabel 1.** Sampel Data Pembelian Pakaian Tahun 2023-2024

ID Customer	Jumlah	Harga Satuan	Total Keseluruhan	Frekuensi Pembelian (1Thn)
001	155	25000	25155	5
002	607	35000	21245000	14
003	56	51000	2856000	1
004	80	70000	6500000	7
005	220	45000	9900000	4
006	300	65000	19500000	5
007	95	70000	6650000	3
008	45	128000	5760000	1
009	904	20000	18080000	4

ID Customer	Jumlah	Harga Satuan	Total Keseluruhan	Frekuensi Pembelian (1Thn)
010	10	135000	1350000	1

Pada tahap ini, data akan dianalisis dan dipilih untuk menentukan variabel yang akan dimasukkan, guna menghindari duplikasi atau pengulangan yang tidak relevan dalam proses data mining. Variabel yang digunakan mencakup customer ID, jumlah pesanan, total harga, dan frekuensi pembelian selama 3 tahun. Tabel 2 adalah dataset yang sudah dipilih.

Tabel 2. Data Selection Pembelian Pakaian

ID Customer	Jumlah	Total Keseluruhan	Frekuensi Pembelian (1Thn)
001	155	25155	5
002	607	21245000	14
003	56	2856000	1
004	80	6500000	7
005	220	9900000	4
006	300	19500000	5
007	95	6650000	3
008	45	5760000	1
009	904	18080000	4
010	10	1350000	1

Pada tahap ini, dilakukan pembersihan data yang tidak lengkap dan validasi data yang tidak konsisten. Data pelanggan akan diperbaiki jika terdapat kesalahan. Variabel yang digunakan adalah variabel yang sama dengan yang digunakan dalam seleksi data. Tabel 3 menunjukkan data yang telah dipreproses.

Tabel 3. Data Preprocessing Atau Data Cleaning

ID Customer	Jumlah	Total Keseluruhan	Frekuensi Pembelian (1Thn)
001	155	25155	5
002	607	21245000	14
003	56	2856000	1
004	80	6500000	7
005	220	9900000	4
006	300	19500000	5
007	95	6650000	3
008	45	5760000	1
009	904	18080000	4
010	10	1350000	1

Transformasi data melibatkan mengubah bentuk data agar sesuai dengan format yang diperlukan oleh algoritma data mining yang akan digunakan. Penyesuaian data ke format yang sesuai dilakukan pada setiap tahap pengolahan data mining. Setelah menyelesaikan tahap seleksi data, pembersihan data, dan transformasi data, langkah berikutnya adalah menerapkan Algoritma K-Means Clustering. Pada tahap ini, data akan dikelompokkan menggunakan Algoritma K-Means Clustering dengan memanfaatkan Visual Studio Code dan bahasa pemrograman Python sebagai alatnya, dengan jumlah data sebanyak 661 pembelian Pakaian dari tahun 2023-2024. Berikut adalah langkah-langkahnya:

Menggunakan Perhitungan Manual di Ms Excel, langkah awal adalah menyiapkan sampel data yang telah divalidasi, menggunakan variabel seperti id pelanggan, jumlah pesanan, total harga, dan frekuensi pembelian selama 1 tahun. sesuai dengan yang tertera dalam Tabel 4 berikut:

Tabel 4. Dataset

ID Customer	Jumlah	Total Keseluruhan	Frekuensi Pembelian (1Thn)
001	155	25155	5
002	607	21245000	14
003	56	2856000	1
004	80	6500000	7
005	220	9900000	4
006	300	19500000	5
007	95	6650000	3
008	45	5760000	1
009	904	18080000	4
010	10	1350000	1

Langkah kedua adalah menentukan jumlah kluster dan memilih titik pusat kluster awal secara acak. Pusat kluster atau centroid ditentukan dengan cara meminimalkan jarak antara titik data dengan pusat kluster, sebagaimana dijelaskan dalam Tabel 5 berikut ini:

**Tabel 5.** Data Titik Pusat Cluster Awal

<b>C1</b>	155	25000	25155	5
<b>C2</b>	607	35000	21245000	14
<b>C3</b>	56	51000	2856000	1

Langkah ketiga melibatkan perhitungan jarak antara setiap objek dengan centroid dari masing-masing kluster menggunakan teori Euclidean. Setiap data akan ditempatkan ke dalam kluster yang memiliki jarak paling dekat dengan data tersebut, seperti yang dijelaskan dalam Tabel 6 berikut:

**Tabel 6.** Hasil Perhitungan Jarak Setiap Objek Ke Centroid

ID Customer	C1	C2	C3	Jarak Terdekat
1	0	190978605	2830845	0
10	1324845	258635000	1506000	1324845
2	23945	0	2854786	0
6	19474845	15705000	16644000	15705000
3	2830845	239057000	0	0
4	6474845	103215000	3644000	3644000
5	9874845	113450000	7044000	7044000
7	6624845	160545000	3794000	3794000
8	5734845	201305000	2904000	2904000
9	18054845	31650000	15224000	15224000

Langkah keempat, setelah jarak dari setiap objek ke centroid ditemukan, adalah mengelompokkan data ke dalam cluster berdasarkan jarak terdekatnya, seperti yang ditunjukkan pada tabel 7 di bawah ini:

**Tabel 7.** Kelompok Cluster

ID Customer	Jumlah	Total Keseluruhan	Frekuensi Pembelian (1Thn)	Jarak Terdekat	Cluster
1	155	25155	5	0	1
10	10	1350000	1	1324845	1
2	607	1214	14	0	2
6	300	19500000	5	15705000	2
3	56	2856000	1	0	3
4	80	6500000	7	3644000	3
5	220	9900000	4	7044000	3
7	95	6650000	3	3794000	3
8	45	5760000	1	2904000	3
9	904	18080000	4	15224000	3

Langkah kelima, setelah semua data dikelompokkan, adalah menghitung ulang pusat kluster dengan mengambil nilai rata-rata dari data yang ada dalam kelompok kluster yang sama. Tabel 8 menunjukkan titik pusat cluster baru yang telah dirata-rata.

**Tabel 8.** Titik Pusat Cluster Baru

Centeroid Baru	Jumlah	Total Keseluruhan	Frekuensi Pembelian (1Thn)
C1	83	80000	687578
C2	607	35000	21245000
C3	243	64143	9892286

Langkah keenam adalah melakukan iterasi dengan mengulangi perhitungan rata-rata pusat cluster beberapa kali, kemudian menentukan jarak terdekat dan mengelompokkan data ke dalam cluster hingga hasilnya sama dengan kelompok cluster awal. Langkah ketujuh, hasil dari cluster yang stabil akan ditampilkan untuk menunjukkan label cluster yang dimiliki oleh anggota cluster tertentu. Hasil cluster yang sama seperti cluster awal ditunjukkan pada tabel 9 berikut:

**Tabel 9.** Data Hasil Cluster Yang Stabil Atau Tetap

ID Customer	Jumlah	Total Keseluruhan	Frekuensi Pembelian (1Thn)	Jarak Terdekat	Cluster Baru	Cluster Lama
1	155	25155	5	0	1	1
10	10	1350000	1	1324845	1	1
2	607	1214	14	0	2	2

ID Customer	Jumlah	Total Keseluruhan	Frekuensi Pembelian (1Thn)	Jarak Terdekat	Cluster Baru	Cluster Lama
6	300	19500000	5	15705000	2	2
3	56	2856000	1	0	3	3
4	80	6500000	7	3644000	3	3
5	220	9900000	4	7044000	3	3
7	95	6650000	3	3794000	3	3
8	45	5760000	1	2904000	3	3
9	904	18080000	4	15224000	3	3

Langkah kedelapan adalah mengevaluasi akurasi Davies Bouldin Index (DBI) dengan memanfaatkan rumus Sum of Square Within cluster (SSW) pertama. Evaluasi ini bertujuan untuk mengukur seberapa erat anggota cluster terikat dalam satu cluster, sebagaimana ditunjukkan pada tabel 10 di bawah ini:

**Tabel 10.** Evaluasi Square Within cluster

C1	C2	C3	JARAK TERDEKAT	CLUSTER	SSW	
0	190978605	2830845	0	1	662423	SSWC1
1324845	258635000	1506000	1324845	1		
23945	0	2854786	0	2	7852500	SSWC2
19474845	15705000	16644000	15705000	2		
2830845	239057000	0	0	3		
6474845	103215000	3644000	3644000	3		
9874845	113450000	7044000	7044000	3	5435000	SSWC3
6624845	160545000	3794000	3794000	3		
5734845	201305000	2904000	2904000	3		
18054845	31650000	15224000	15224000	3		

Langkah kesembilan adalah mengevaluasi akurasi dengan Davies Bouldin Index (DBI) menggunakan rumus Sum of Square Between Cluster (SSB). Rumus ini digunakan untuk mengukur separasi atau perbedaan antara satu cluster dengan cluster lainnya, seperti yang ditunjukkan pada tabel 11 di bawah ini:

**Tabel 11.** Evaluasi Square Between Cluster

SSB	C1	C2	C3
C1	0	689756	9892358
C2	689756	0	23377035
C3	9892358	23377035	0

Langkah kesebelas adalah mengevaluasi akurasi Davies Bouldin Index (DBI) dengan menggunakan rumus ketiga rasio untuk menentukan seberapa baik nilai perbandingan antara satu cluster dengan cluster lainnya, seperti yang ditunjukkan pada tabel 12 di bawah ini.

**Tabel 12.** Hasil Akurasi Davies Bouldin Index

R1,2	20
R1,3	1
R2,3	1

Langkah duabelas dalam evaluasi akurasi menggunakan Davies Bouldin Index (DBI) melibatkan penggunaan rumus keempat, yaitu Davies Bouldin Index (DBI) itu sendiri. DBI dianggap semakin baik ketika nilainya mendekati nol tetapi tidak negatif. Hasil dari DBI ditunjukkan dalam tabel 13 di bawah ini:

**Tabel 13.** Hasil Davies Bouldin Index (DBI)

DBI
10
1
0

#### 4. KESIMPULAN

Penelitian ini menerapkan Algoritma K-Means Clustering untuk mengelompokkan tingkat loyalitas pelanggan di PT.Bhara Utama, serta mengevaluasi keakuratannya dengan Davies-Bouldin Index (DBI). Eksperimen dilakukan pada 660 data pelanggan, yang menghasilkan pembagian menjadi 3 kelompok: 8 pelanggan dalam kelompok yang sangat loyal, 461 pelanggan dalam kelompok cukup loyal, dan 191 pelanggan dalam kelompok tidak loyal. Evaluasi akurasi

menggunakan DBI menunjukkan nilai sebesar 0,19. Semakin kecil nilai DBI, semakin baik kualitas kluster yang terbentuk.

## REFERENCES

- [1] A. Adiwijaya, M. N. Aulia, M. S. Mubarak, W. U. Novia and F. Nhita, “A comparative study of MFCC-KNN and LPC-KNN for hijaiyyah letters pronunciation classification system,” *Information and Communication Technology (ICoICT)*, pp. (pp. 1-5), 2017.
- [2] I. A. Adriana, *Penalaran Komputer Berbasis Kasus (Case Based Reasoning)*, Yogyakarta: Ardana Media, 2007.
- [3] M. N. Al-Kabi, G. Kanaan, R. Al-Shalabi, S. Al-Sinjalawi dan R. S. Al-Mustafa, “Al-Hadith Text Classifier,” *Journal of Applied Sciences* 5, pp. 584-587, 2005.
- [4] A. Kusumaningrum and S. Al-Faraby , “Klasifikasi Informasi, Anjuran dan Larangan pada Hadits Shahih Bukhari menggunakan Metode Support Vector Machine,” *e-Proceeding of Engineering*, p. 5014, 2017.
- [5] A. I. Pratiwi and Adiwijaya, “On the Feature Selection and Classification Based on Information Gain for Document Sentiment Analysis,” *Applied Computational Intelligence and Soft Computing*, vol. 2018, p. 5, 2018.
- [6] A. Desiani dan M. Arhami, *Konsep Kecerdasan Buatan*, 1 penyunt., D. Hardjono, Penyunt., Yogyakarta: Penerbit ANDI, 2006.
- [7] E. Riviera, S. Al-Faraby and Adiwijaya , “Klasifikasi Anjuran, Larangan dan Informasi pada Hadis Sahih Al-Bukhari” *e-Proceeding of Engineering*, p. 4683, 2017.
- [8] S. W. Faza Akmal, “SISTEM PPAKAR UNTUK MENDIAGNOSA PENYAKIT LAMBUNG DENGAN IMPLEMENTASI METODE CBR (CASE BASED REASONING) BERBASIS WEB,” *Jurnal Sarjana Teknik Informatika*, vol. 2 , no. 1, Februari 2014.
- [9] F. Harrag dan E. El-Qawasmah, “Neural Network for Arabic Text Classification,” *2009 Second International Conference on the Applications of Digital Information and Web Technologies*, pp. 778-783, 2009.
- [10] H. N. A. Hamed, S. M. Shamsuddin and N. Salim. , “Particle Swarm Optimization For Neural Network Learning Enhancement,” *Jurnal Teknologi*, pp. 13-26, 2008.
- [11] Kusrini, *Sistem Pakar Teori dan Aplikasi*, 1 penyunt., f. Suyantoro, Penyunt., Yogyakarta: Penerbit ANDI, 2006.
- [12] M. L. Zhang and Z. H. Zhou , “Multilabel neural networks with applications to functional genomics and text,” *IEEE transactions on Knowledge and Data Engineering*, pp. 1338-1351, 2006.
- [13] S. Nurcahyo, F. Nhita and Adiwijaya, “Rainfall Prediction in Kemayoran Jakarta Using Hybrid Genetic Algorithm (GA) and Partially Connected Feedforward Neural Network (PCFNN),” *Information and Communication Technology (ICoICT)*, pp. (pp. 166-171), 2014.
- [14] T. Sutojo, E. Mulyanto and V. Suhartono, *Kecerdasan Buatan*, Yogyakarta: Andi, 2011.
- [15] E. A. Soepardi, N. Iskandar, J. Bashiruddin, R. D. Restuti and H. Utama, *Telinga Hidung Tenggorok Kepala dan Leher edisi ketujuh*, Jakarta: FK UI, 2012.
- [16] U. KEKEVÍ, “Real-Time Big Data Processing and Analytics: Concepts, Technologies, and Domains,” *Computer Science*, no. 2, pp. 111-123, 2022.
- [17] R. A. Pane, M. S. Mubarak, N. S. Huda and Adiwijaya, “A Multi-lable Classification on Topics of Quranic Verses in English Translation using Multinomial Naive Bayes,” *6th International Conference on Information and Communication Technology (ICoICT)*, 2018.
- [18] J. S. D. Raharjo, “Model Artificial Neural Network berbasis Particle Swarm Optimization untuk Prediksi Laju Inflasi,” *Sistem Komputer*, 2013.
- [19] N. H. Harani, “Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python,” *Jurnal Manajemen Informatika (JAMIKA)*, vol. 10, no. 2, pp. 133-146, 2020.
- [20] Yulia, “Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means,” *Indonesian Journal of Computer Science*, vol. 10, no. 1, 2021.