

Pengklasteran dan Segmentasi Karakteristik Donatur Sedekah Daring Dengan Teknik Penambahan Data

Martono¹, Mohammad Syafrullah²

^{1,2}Universitas Budi Luhur, Jl. Ciledug Raya, Kec. Pesanggrahan, Kota Jakarta Selatan, Indonesia
E-mail: tono_110385@yahoo.com¹, mohammad.syafrullah@budiluhur.ac.id²

Abstract – PPPA Daarul Qur'an is actively involved in collecting charity funds through various digital platforms. They focus on promoting self-reliance in communities through programs such as Qur'an memorization. However, they face challenges in identifying potential donors and targeting their fundraising campaigns effectively. In this study, donor data segmentation was done using FCM algorithm for similarity-based clustering and DBSCAN for density-based clustering. The evaluation results showed that DBSCAN provided better clustering results compared to FCM, with higher values for Calinski-Harabasz Index and Silhouette Score, indicating better cluster separation and homogeneity. The use of LRFM variables illustrated donor behavior during donations. The evaluation values for FCM were SSE: 4563140439.7347, CHI: 0.3128562606910544, SS: 0.08061440564597909, and DBI: 9.910899528742423. For DBSCAN, the values were CHI: 3237127.1389106703, SS: 0.8479515063332151, and DBI: 2.1939426200975047. Based on the evaluation, DBSCAN seemed to be more effective for this dataset in terms of cluster separation and homogeneity.

Keywords - LRFM, DBSCAN, FCM, *Clustering*, CLV

Intisari – PPPA Daarul Qur'an terlibat aktif dalam penggalangan dana amal melalui berbagai platform digital. Mereka fokus pada peningkatan kemandirian masyarakat melalui program seperti penghafalan Al-Qur'an. Namun, mereka menghadapi tantangan dalam mengidentifikasi donor potensial dan menargetkan kampanye penggalangan dana secara efektif. Pada penelitian ini, segmentasi data donatur dilakukan dengan menggunakan algoritma FCM untuk clustering berbasis kesamaan dan DBSCAN untuk clustering berbasis kepadatan. Hasil evaluasi menunjukkan bahwa DBSCAN memberikan hasil clustering yang lebih baik dibandingkan dengan FCM, dengan nilai Calinski-Harabasz Index dan Silhouette Score yang lebih tinggi, menunjukkan pemisahan cluster dan homogenitas yang lebih baik. Penggunaan variabel LRFM menggambarkan perilaku donatur selama berdonasi. Nilai evaluasi FCM adalah SSE: 4563140439.7347, CHI: 0.3128562606910544, SS: 0.08061440564597909, dan DBI: 9.910899528742423. Untuk DBSCAN nilainya adalah CHI: 3237127.1389106703, SS: 0.8479515063332151, dan DBI: 2.1939426200975047. Berdasarkan evaluasi, DBSCAN tampaknya lebih efektif untuk kumpulan data ini dalam hal pemisahan cluster dan homogenitas.

Kata Kunci - LRFM, DBSCAN, FCM, *Clustering*, CLV

I. PENDAHULUAN

PPPA Daarul Qur'an aktif dalam mengumpulkan dana amal seperti sedekah, zakat, dan wakaf secara online, yang menjadi tambang data berharga di era digital [1]. Organisasi ini mendukung program dakwah, tahfizhul Qur'an, bantuan kemanusiaan, beasiswa, kesehatan, dan pengembangan masyarakat. Analisis segmentasi donatur penting untuk meningkatkan pendapatan donasi dan memastikan penyaluran yang tepat sasaran, menjaga kepercayaan masyarakat sebagai pengelola ziswaf untuk pembangunan bangsa. Kepuasan dan loyalitas donatur krusial dalam hubungan jangka panjang dengan organisasi [2]. Laznas PPPA Daarul Qur'an telah mengelompokkan donatur menjadi tiga segmen berdasarkan jumlah donasi.

Oleh karena itu, model RFM (*Recency, Frequency, Monetary*) diusulkan untuk mengklasifikasikan donatur secara lebih komprehensif. Model RFM, meskipun populer, memiliki keterbatasan dalam mengidentifikasi loyalitas donatur. Sebagai solusi, model LRFM (*Length, Recency, Frequency, Monetary*) dikembangkan untuk mengukur loyalitas donatur dengan lebih baik. Penelitian ini juga menggunakan teknik penambangan data dengan algoritma FCM (*Fuzzy C-Means*) dan DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) untuk mengelompokkan donatur berdasarkan pola interaksi mereka. Penggunaan CRM (*Customer Relationship Management*) penting dalam memproses interaksi antara donatur dan lembaga, dengan teknik klasifikasi dan klastering untuk memahami profil donatur dan merancang strategi interaksi yang efektif [3]. Dalam mengelola data donatur yang kompleks, penting untuk memilih atribut yang tepat agar analisis data optimal dan bermakna. Model LRFM membantu mengidentifikasi atribut relevan seperti interval waktu awal donasi (*Length*), jarak waktu terakhir donasi (*Recency*), frekuensi donasi (*Frequency*), dan total donasi (*Monetary*).

Penelitian ini mengintegrasikan temuan dan metodologi dari studi-studi terdahulu yang relevan dalam bidang pengklasteran dan segmentasi karakteristik donatur sedekah daring. Penelitian sebelumnya menunjukkan bahwa algoritma pengklasteran seperti K-Means dan *Fuzzy C-Means* efektif dalam membagi populasi pengguna daring berdasarkan perilaku atau karakteristik tertentu, yang sesuai dengan karakteristik donatur sedekah online yang bervariasi dalam tingkat kontribusi, frekuensi donasi, dan preferensi platform (X, Y). Integrasi teknik data mining juga diperkuat oleh penelitian sebelumnya yang menunjukkan bahwa analisis data ini dapat memberikan wawasan yang mendalam tentang perilaku donatur online, membantu dalam pengembangan strategi pemasaran yang lebih personal, serta peningkatan retensi donatur (A, B).

Penelitian ini mengusulkan model LRFM dan algoritma FCM serta DBSCAN untuk segmentasi donatur yang efisien dan objektif. Algoritma FCM dipilih karena kemampuannya untuk menangani data yang tidak terstruktur dengan baik dan mengizinkan satu data masuk ke dalam lebih dari satu cluster, yang memungkinkan analisis yang lebih fleksibel dan realistis. Sementara itu, DBSCAN dipilih karena kemampuannya dalam mengidentifikasi cluster dengan bentuk yang arbitrer dan menemukan pola data yang mengandung noise, sehingga cocok untuk data donatur yang memiliki variasi besar. Kedua algoritma ini, ketika digabungkan, memberikan pendekatan yang lebih komprehensif dan akurat dalam mengelompokkan donatur berdasarkan pola interaksi mereka.

Penelitian ini bertujuan untuk mengidentifikasi pola atau kelompok-kelompok karakteristik donatur sedekah daring menggunakan teknik penambangan data. Hipotesis yang ingin diuji dalam penelitian ini adalah apakah terdapat pola atau kelompok-kelompok karakteristik donatur sedekah daring yang dapat diidentifikasi berdasarkan data yang dikumpulkan. Selain itu penelitian ini akan mencoba untuk mengklaster dan mengsegmentasi donatur sedekah daring berdasarkan berbagai variabel seperti usia, jenis kelamin, pendapatan, frekuensi donasi, dan lain sebagainya. Melakukan identifikasi pola-pola tersebut, diharapkan dapat memberikan wawasan yang lebih dalam tentang karakteristik donatur sedekah daring sehingga dapat membantu organisasi atau lembaga yang bergerak di bidang sosial untuk lebih efektif dalam merancang program-program penggalangan dana atau kegiatan sosial lainnya.

Penelitian mengenai pengklasteran dan segmentasi karakteristik donatur sedekah daring dengan teknik penambangan data sangat penting dilakukan karena dapat memberikan wawasan yang mendalam mengenai perilaku dan preferensi donatur dalam melakukan sedekah secara daring. Memiliki pemahaman yang lebih baik mengenai karakteristik donatur, organisasi dan donor dapat meningkatkan efektivitas dan efisiensi program sedekah yang mereka jalankan. Adanya metode segmentasi yang lebih baik, organisasi dapat menyesuaikan strategi pemasaran dan komunikasi mereka untuk mencapai target donatur dengan lebih tepat dan efisien. Hal ini

diharapkan dapat meningkatkan partisipasi donatur, memperluas jangkauan program sedekah, serta meningkatkan keberlangsungan program secara keseluruhan. Oleh karena itu, penelitian ini diharapkan dapat memberikan dampak positif bagi organisasi dalam meningkatkan efektivitas program sedekah mereka dan juga bagi donor dalam memberikan kontribusi yang lebih bermakna.

Penelitian ini memiliki implikasi yang signifikan baik dalam konteks praktis maupun akademis. Secara praktis, hasil penelitian ini dapat meningkatkan efisiensi program sedekah daring dengan mengidentifikasi dan mengklaster donatur berdasarkan karakteristik mereka. Hal ini dapat membantu lembaga atau platform sedekah untuk menargetkan donatur potensial dengan lebih tepat, meningkatkan partisipasi dalam program sedekah, serta meningkatkan keberlanjutan program dengan strategi yang lebih efektif. Selain itu, pengetahuan yang diperoleh dari penelitian ini juga dapat digunakan untuk mengembangkan strategi komunikasi yang lebih efektif dan berkelanjutan. Secara akademis, penelitian ini dapat memperluas pemahaman tentang teori segmentasi donatur dalam konteks sedekah daring, serta mengembangkan metode dan teknik pengklasteran yang dapat diterapkan dalam penelitian lanjutan dalam bidang ini. Oleh karena itu, penelitian ini tidak hanya memberikan nilai tambah dalam praktik langsung bagi lembaga atau platform sedekah, tetapi juga dalam pengembangan teori dan metode akademis yang relevan dalam konteks sosial dan pengklasteran data.

II. SIGNIFIKANSI STUDI

A. Penelitian Terdahulu

Investigasi ilmiah sebelumnya, yang merupakan landasan dasar untuk penyelidikan ini, digambarkan dalam Tabel 1.

TABEL I
PENELITIAN TERDAHULU

Nomor	Penulis	Penelitian Terdahulu
1	Sudriyanto [4]	<i>Clustering</i> Loyalitas Pelanggan Dengan Metode RFM (<i>Recency, Frequency, Monetary</i>) dan <i>Fuzzy C-Means</i> Penelitian ini menunjukkan keunggulan metode <i>Fuzzy C-Means</i> (FCM) dalam clustering data. Dengan eksponen pangkat 2, FCM menghasilkan <i>Partition Coefficient Index</i> (PCI) sebesar 0,8156 dan <i>Indeks Xie Beni</i> (XBI) sebesar 0,0069, menunjukkan clustering yang baik dengan minimal overlapping. Eksponen pangkat 3 menghasilkan PCI 0,5860 dan XBI 0,0632. FCM fleksibel dalam penyesuaian eksponen pangkat untuk hasil optimal. Kemampuan FCM menangani data tidak terstruktur memungkinkan satu data menjadi anggota lebih dari satu cluster, memberikan kejelasan pada batas-batas cluster yang tumpang tindih, bermanfaat untuk analisis data kompleks seperti karakteristik donatur sedekah daring.
2	Denny B. Saputra [5]	Implementasi <i>Fuzzy C-Means</i> dan Model RFM untuk Segmentasi Pelanggan (Studi Kasus : PT. XYZ) Metode penelitian yang menggunakan <i>Analytic Hierarchy Process</i> (AHP) dan <i>Elbow</i> memiliki beberapa kelebihan. AHP membantu mengidentifikasi variabel moneter sebagai yang paling penting dengan bobot 0,59, diikuti oleh variabel kekinian (0,31) dan frekuensi (0,10). Ini memungkinkan prioritas yang jelas dalam analisis data. Penggunaan metode <i>Elbow</i> untuk menentukan jumlah cluster optimal, yaitu $K = 3$, memastikan segmentasi yang efisien. Hasilnya, 123 pelanggan tersegmentasi menjadi tiga cluster dengan distribusi yang tepat: 42 pelanggan di <i>cluster</i> 1, 64 di <i>cluster</i> 2, dan 17 di <i>cluster</i> 3, memberikan dasar yang kuat untuk strategi pemasaran yang lebih tepat sasaran.
3	Alfian [6]	<i>Clustering</i> Kesetiaan Pelanggan E-Ritel Dengan Model RFM (<i>Recency, Frequency, Monetary</i>) Dan <i>K-Means</i> Hasil penelitian menunjukkan bahwa pendekatan RFM (<i>Recency, Frequency, Monetary</i>) menghasilkan cluster dengan interpretasi sebagai berikut: peningkatan nilai <i>Recency</i> (R) menandakan seringnya transaksi pelanggan; nilai <i>Frekuensi</i> (F)

Nomor	Penulis	Penelitian Terdahulu
		yang tinggi menunjukkan loyalitas pelanggan yang kuat; dan peningkatan nilai <i>Moneter</i> (M) mencerminkan jumlah transaksi yang besar. Pemanfaatan <i>Indeks Davies-Bouldin</i> (DBI) memastikan ketepatan dan keakuratan cluster yang digunakan, meningkatkan keandalan hasil analisis. Pendekatan RFM dalam penelitian ini memiliki beberapa kelebihan, membantu memahami perilaku pelanggan serta memastikan hasil yang valid dan dapat diandalkan.
4	Nindya Rahmawati Syarif [7]	Implementasi <i>Algoritma Fuzzy CMeans</i> Dan Metode <i>Recency Frequency Monetary</i> (RFM) Pada Aplikasi Data Mining Untuk Pengelompokan Pelanggan Metode <i>Fuzzy C-Means</i> (FCM) yang digunakan dalam penelitian ini sangat efektif dalam segmentasi pelanggan berdasarkan atribut RFM (<i>Recency, Frequency, Monetary</i>). FCM berhasil mengelompokkan pelanggan PT Eka Cipta Rasa menjadi empat kategori: <i>Golden</i> (38 individu), <i>Silver</i> (186 individu), <i>Bronze</i> (103 individu), dan <i>Iron</i> (0 individu). Algoritma FCM mencapai nilai fungsi tujuan kumulatif sebesar 102,2 dengan perubahan fungsi tujuan sebesar 51,1, dalam 9 iterasi. Keunggulan FCM terletak pada kemampuannya dalam mengelompokkan data tidak terstruktur dengan baik, menghasilkan clustering yang jelas, yang sangat penting untuk strategi pemasaran yang efektif.
5	Andri Syafrianto [8]	Perancangan Aplikasi K-Means Untuk Pengelompokan Mahasiswa STMIK ELRAHMA Yogyakarta Berdasarkan Frekuensi Kunjungan KePerpustakaan dan IPK Temuan penelitian ini mengindikasikan bahwa mahasiswa dengan IPK rendah dan jarang berkunjung ke perpustakaan cenderung berpusat pada koordinat (1.48; 1.4). Sebaliknya, mahasiswa dengan IPK tinggi dan sering berkunjung ke perpustakaan cenderung berkumpul di sekitar pusat yang berada pada koordinat (3.35; 1.64). Selain itu, mahasiswa dengan IPK tinggi dan frekuensi kunjungan ke perpustakaan yang tinggi juga menunjukkan pusat cluster pada koordinat (3.49; 3.5). Metode penelitian ini memiliki beberapa kelebihan. Pertama, memberikan gambaran yang jelas tentang hubungan antara IPK, kunjungan ke perpustakaan, dan pola cluster. Kedua, memungkinkan identifikasi visual pusat-pusat cluster melalui koordinat yang dihasilkan. Ketiga, menggunakan pendekatan analisis spasial yang kuat untuk menggambarkan distribusi spasial dari pola yang diamati. Keempat, memberikan informasi bermanfaat untuk pengambilan keputusan dan pengembangan strategi dalam meningkatkan kinerja akademik mahasiswa.

B. Tinjauan Pustaka

1. Huruf-huruf Dokumen

Penambangan data adalah proses berulang untuk menemukan wawasan baru dengan mengidentifikasi struktur rumit atau pola berulang dalam kumpulan data yang luas [11]. Tujuan dari tahap eksplorasi ini adalah untuk mengungkap pola tersembunyi dalam kumpulan data besar yang disimpan di database, gudang data, atau sistem penyimpanan data alternatif [12]. Data mining juga dikenal sebagai KDD (*Knowledge Discovery in Database*), yang melibatkan pencarian dan pemanfaatan data histori transaksi untuk mengeksplorasi pola, keteraturan, atau hubungan dalam kumpulan data besar.

2. Cross Industry Standard Process for Data Mining (CRISP DM)

Kegiatan penggalian data (data mining) merupakan komponen penting dalam tahapan KDD, terutama terkait dengan identifikasi dan analisis pola. CRISP-DM adalah kerangka proses yang digunakan dalam data mining dan berfungsi sebagai dasar strategi pemecahan masalah [13].

3. Length, Recency, Frequency, Monetary (Pemodelan LRFM)

RFM merupakan singkatan dari *Recency, Frequency, dan Monetary* yang mana metode ini sering digunakan untuk menilai nilai donatur dengan mempertimbangkan waktu terakhir donasi (*recency*), frekuensi donasi (*frequency*), dan jumlah total donasi (*monetary*) [14]. Metode ini efektif dalam melakukan segmentasi donatur dengan membagi mereka ke dalam kelompok-kelompok tertentu, sehingga layanan personal dapat diberikan di masa depan dan donatur yang potensial untuk merespons promosi dapat diidentifikasi. RFM telah diterapkan dalam berbagai

bidang, termasuk industri elektronik dan keamanan komputer. Modifikasi model RFM yang dikenal sebagai LRFM memperkenalkan variabel "Panjang", yang mewakili durasi dari donasi awal hingga donasi terbaru. Peningkatan ini mengatasi keterbatasan model RFM, yang gagal membedakan donor dengan riwayat kontribusi yang berbeda-beda [15]. Loyalitas donatur sangat bergantung pada periode interaksi antara donatur dan lembaga, sehingga diperlukan tambahan variabel untuk mengukur tingkat loyalitas yang telah dibangun oleh donatur. Oleh karena itu, beberapa penelitian lebih cenderung mengadopsi model LRFM untuk melakukan segmentasi terhadap donatur suatu lembaga.

4. *Customer Relationship Management (CRM)*

CRM (*Customer Relationship Management*) adalah mewakili pendekatan holistik yang bertujuan untuk membangun, memelihara, dan memperluas interaksi dengan pelanggan dari waktu ke waktu [16]. Proses CRM menyediakan kerangka kerja untuk membangun hubungan yang sukses dengan donatur. Selama beberapa dekade terakhir, peran penting konsumen dalam mengarahkan strategi lembaga atau perusahaan semakin diakui. CRM mencakup berbagai upaya untuk memahami donatur dan memanfaatkan pengetahuan ini dalam merancang dan melaksanakan kegiatan kampanye ziswaf [17].

5. *Density Based Spatial Clustering Application with Noise (DBSCAN)*

DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) adalah algoritma *clustering* yang mengelompokkan data berdasarkan kerapatan spasial. Algoritma ini mampu mengidentifikasi kelompok data dengan kepadatan tinggi (*core points*), data yang berada di sekitar titik inti (*border points*), dan data yang berada di area dengan kepadatan rendah (*noise points*) [18].

6. *Fuzzy C-Means (FCM)*

FCM, yang dikembangkan oleh Jim Bezdek pada tahun 1981, adalah metode klusterisasi berbasis logika *fuzzy* yang berfungsi untuk mengelompokkan data ke dalam kelompok-kelompok yang sesuai berdasarkan nilai keanggotaan *fuzzy* masing-masing data. Dalam FCM, setiap elemen data memiliki derajat keanggotaan dalam setiap kluster, menunjukkan sejauh mana data tersebut termasuk dalam kluster tertentu [19]. Pada tahap awal, FCM menentukan pusat kluster yang mewakili posisi rata-rata dari setiap kluster. Meskipun posisi pusat kluster awal belum sepenuhnya akurat, melalui proses iterasi, posisi ini akan dihitung ulang dan menjadi lebih tepat. Hal yang sama berlaku untuk nilai keanggotaan setiap data, yang akan disesuaikan secara bertahap hingga mencapai akurasi yang lebih tinggi [20].

7. *Clustering*

Clustering adalah teknik umum yang digunakan untuk mengelompokkan data ke dalam kelompok-kelompok (*cluster*) sehingga setiap kelompok memiliki data yang mirip satu sama lain dan berbeda dengan kelompok lainnya [21]. Dalam *clustering* berbasis jarak, kita mencari sejumlah pusat *cluster* di dalam himpunan data X yang dapat merepresentasikan ciri khas dari masing-masing *cluster* di X. Kriteria kesamaan dalam hal ini diukur dengan jarak geometris. Di sisi lain, pendekatan *conceptual clustering* mengelompokkan objek berdasarkan sejauh mana objek tersebut cocok dengan konsep deskriptif.

8. *Segmentasi Donatur*

Segmentasi merupakan alat analisis yang vital dalam merancang strategi efektif dan dengan segmentasi, donatur dapat dibagi menjadi beberapa segmen berdasarkan kriteria tertentu untuk mempersiapkan strategi kampanye donasi. Langkah awal dalam pemodelan data donatur adalah melakukan segmentasi donasi [22].

9. *Analytical Hierarchy Process (AHP)*

Analytical Hierarchy Process (AHP) diperkenalkan oleh Thomas L. Saaty pada tahun 1970-an, merupakan metode pengambilan keputusan multi-kriteria yang populer dan telah berhasil diterapkan dalam berbagai konteks pengambilan keputusan praktis [23]. AHP

memungkinkan permasalahan kompleks dibagi ke dalam hierarki berdasarkan atribut atau kriteria yang relevan.

10. *Metode Elbow*

Metode Elbow digunakan untuk menentukan jumlah *cluster* yang optimal dengan melihat perbandingan antara jumlah *cluster* dan sudut siku yang terbentuk pada suatu titik. Persentase hasil dari setiap hitungan direpresentasikan dalam bentuk grafik untuk dibandingkan. Jika terdapat penurunan yang signifikan pada nilai *cluster* dari *cluster* pertama ke *cluster* kedua, maka jumlah *cluster* tersebut dianggap sebagai pilihan terbaik [6].

11. *Sum of Square Error (SSE)*

Sum of Square Error (SSE) adalah salah satu metode sederhana yang digunakan untuk mengevaluasi hasil dari proses pengelompokan. Dalam perhitungan SSE, setiap objek dalam sebuah kelompok akan diukur tingkat kesalahannya terhadap centroid terdekat. Semakin kecil nilai SSE, semakin baik hasil pengelompokan yang diperoleh [19].

12. *Calinski-Harabasz Index (CH)*

CH atau *Criteria Variance Ratio*, dikenal sebagai metrik yang digunakan untuk mengevaluasi model pengelompokan. CH mengukur rasio antara dispersi antar kluster dan dispersi dalam kluster, dihitung sebagai jumlah kuadrat jarak antara titik-titik data [10].

13. *Silhouette Coefficient (SC)*

Silhouette Coefficient (SC) adalah metrik evaluasi untuk mengukur kualitas dari hasil pengelompokan (*cluster*). SC berkisar dari -1 hingga 1, di mana nilai -1 menunjukkan pengelompokan yang salah, nilai 1 menunjukkan pengelompokan yang sangat padat, dan nilai sekitar 0 mengindikasikan adanya tumpang tindih antar *cluster*. Nilai SC yang lebih tinggi menunjukkan bahwa model memiliki *cluster* yang lebih jelas atau memiliki performa yang lebih baik [18].

14. *Davies Bouldin Index (DB)*

Nilai DB adalah ukuran rata-rata "kesamaan" antara *cluster*, yang membandingkan jarak antar *cluster* dengan ukuran *cluster* itu sendiri [12].

15. *Customer Lifetime Value (CLV)*

Customer Lifetime Value (CLV) merupakan ukuran numerik dari total aliran kas bersih yang diperoleh dari seorang donatur selama periode hubungannya dengan sebuah organisasi [11]. Pengukuran ini memiliki nilai penting dan banyak digunakan oleh berbagai lembaga, seperti lembaga keuangan, toko ritel, perusahaan telekomunikasi, dan lain-lain. CLV membantu organisasi untuk membedakan antara donatur mereka, sehingga mereka dapat menyediakan layanan yang sesuai untuk setiap donatur. Dari perhitungan CLV yang akurat, lembaga dapat dengan mudah mengelompokkan donatur ke dalam berbagai kelompok, sehingga strategi kampanye yang tepat dapat dikembangkan untuk setiap kelompok donatur.

C. *Metode Penelitian*

1. *Metode Penelitian*

Metodologi penelitian ini dimulai dengan mengidentifikasi permasalahan, membatasi lingkup, menetapkan tujuan, dan implikasi penelitian. Studi literatur dilakukan untuk referensi yang relevan [9]. Data donasi donatur dikumpulkan dan dibersihkan. Data dibagi menjadi variabel *length*, *recency*, *frequency*, dan *monetary (LRFM)* untuk model. Normalisasi data LRFM dilakukan untuk seragamkan skala nilainya. Bobot LRFM ditentukan dengan metode AHP. *Customer Lifetime Value (CLV)* digunakan untuk mengukur nilai donatur. CLV membantu menentukan segmen fokus. Data LRFM dinormalisasi dan diakumulasi dengan bobot, menghasilkan nilai LRFM bobot untuk clustering. Jumlah *cluster (K)* ditentukan dengan metode *Elbow*, dan proses clustering dilakukan terhadap data LRFM bobot. Hasil clustering divisualisasikan dan karakteristik segmen dianalisis secara mendalam

Namun, ada beberapa kelemahan yang perlu diperhatikan. *Algoritma Fuzzy C-Means* sensitif terhadap inialisasi titik-titik pusat klaster awal, yang dapat menghasilkan solusi yang tidak optimal. Selain itu, hasilnya sering memerlukan interpretasi yang lebih dalam dan kontekstual untuk diterapkan dalam pengambilan keputusan. Pemilihan parameter seperti jumlah klaster dan parameter fuzziness juga krusial, karena pemilihan yang salah dapat mempengaruhi hasil klaster yang dihasilkan. Untuk mengatasi kelemahan ini, disarankan untuk melakukan validasi hasil klaster dengan metode lain, mengoptimalkan parameter algoritma, dan memberikan pemahaman kontekstual yang lebih dalam. Oalah karena itu, penelitian ini tidak hanya memberikan kontribusi teoritis dalam klustering dan segmentasi, tetapi juga dapat memberikan wawasan praktis yang digunakan dalam pengambilan keputusan strategis terkait sedekah daring.

2. Metode Pemilihan Populasi dan Sample

Metodologi pemilihan sampel yang digunakan dalam penelitian ini meliputi tinjauan pustaka yang berfokus pada subjek seperti pengumpulan data, pengelompokan, *Fuzzy C-Means* (FCM), DBSCAN, metode LRFM, dan segmentasi donor, semuanya berkaitan dengan konteks penelitian. Proses tinjauan literatur melibatkan sumber informasi dari berbagai sumber, termasuk wawancara dengan para ahli, referensi dari teks ilmiah, dan dokumen terkait yang berkaitan dengan tema penelitian.

3. Metode Pengumpulan Data

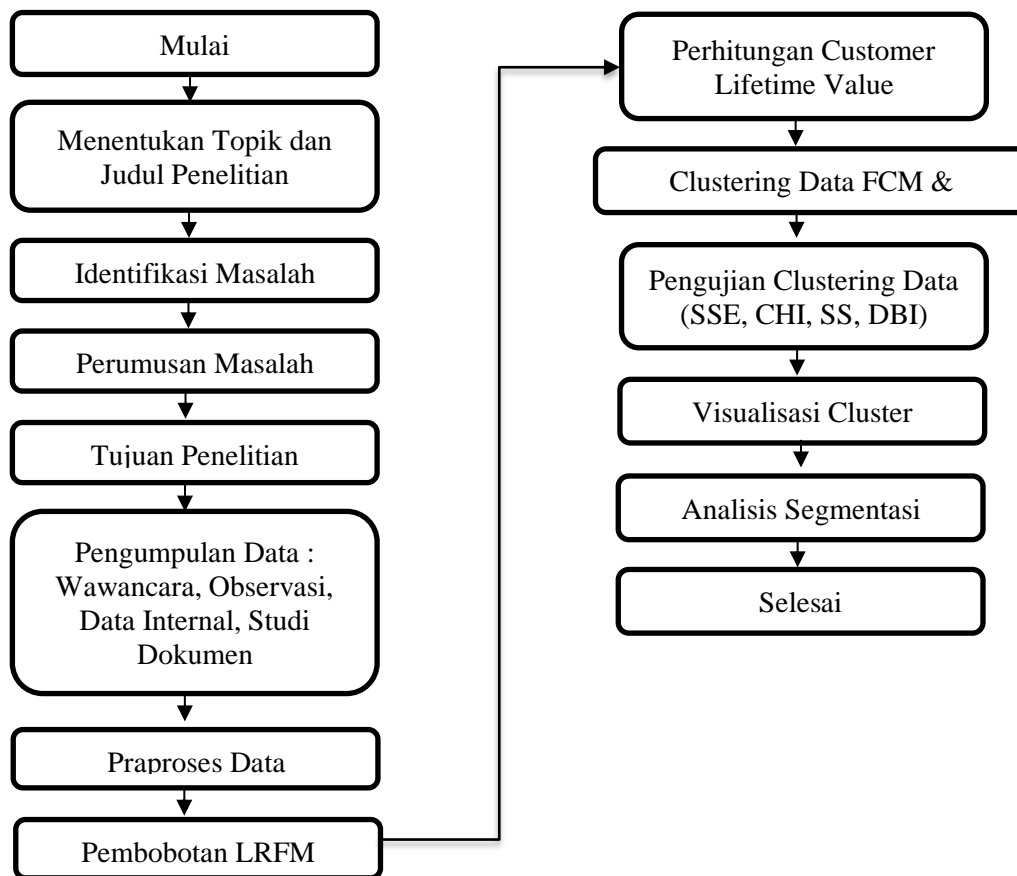
Pada penelitian ini, metode pengumpulan data dilakukan dengan fokus utama pada teknik pengelompokan data (*clustering*). Terdapat beberapa metode yang digunakan, antara lain wawancara, pengamatan, dan penggunaan data internal. Wawancara dilakukan dengan narasumber seperti Direktur Marketing Komunikasi, Ibu Dwi Kartika, serta penanggung jawab dari divisi layanan CRM, pppa.id, dan sedekahonline.com. Wawancara ini bertujuan untuk mengumpulkan informasi tentang proses penerimaan dan penyaluran donasi, serta kriteria-kriteria yang digunakan dalam strategi pemasaran untuk mengelompokkan data donatur. Pengamatan dilakukan di Laznas PPPA Daarul Qur'an Tangerang Banten untuk memahami proses tersebut secara langsung. Selain wawancara dan pengamatan, data internal juga digunakan dalam penelitian ini. Data yang dikumpulkan mencakup informasi tentang donasi donatur, riwayat donasi, tanggal dan waktu donasi, jumlah donasi, serta data lain yang relevan untuk analisis *clustering*. Penggunaan metode ini penting untuk memahami pola interaksi donatur dengan lembaga, serta untuk merancang strategi pemasaran yang lebih efektif.

4. Analisis Data

Teknik analisis deskriptif digunakan untuk menganalisis data yang diperoleh dari studi pustaka, wawancara, dan observasi. Analisis ini bertujuan untuk mendapatkan spesifikasi kebutuhan segmentasi data donatur. Langkah pertama adalah menseleksi dataset dari sumber data, kemudian melakukan seleksi fitur menggunakan model LRFM, dan menyiapkan dataset yang telah difiksasi. Selanjutnya, dilakukan klasterisasi dengan menggunakan Algoritma FCM (*Fuzzy C-Means*) dan DBSCAN, serta evaluasi menggunakan SSE (*Sum Square Error*), *Calinski-Harabasz Index* (CH), *Silhouette Coefficient*, dan *Davies Bouldin Index* (DBI). Untuk aplikasi klasterisasi, peneliti menggunakan platform Streamlit. Proses pengolahan data donatur meliputi beberapa langkah, yakni membersihkan data untuk menghapus baris data kosong atau tidak valid, memilih data berdasarkan atribut seperti jangka waktu donasi, frekuensi donasi, dan nominal donasi dengan model LRFM, dan mempersiapkan data dengan mengurangi kolom yang tidak diperlukan untuk klasterisasi. Data kemudian diubah menjadi format yang terukur agar dapat digunakan sebagai atribut dalam aplikasi Streamlit untuk klasterisasi.

5. Langkah-Langkah Penelitian

Berikut adalah langkah-langkah penelitian yang digunakan sebagai panduan dalam menyusun proses penelitian agar terstruktur dan sistematis.



Gambar 1. Diagram Alur Penelitian

Penelitian ini dimulai dengan menentukan topik dan judul penelitian. Tahap ini merupakan proses pemilihan pokok penelitian yang akan difokuskan pada bidang ilmu tertentu. Topik penelitian menetapkan cakupan kajian yang membedakannya dari penelitian lain, sedangkan judul penelitian merupakan perspektif khusus dari suatu topik. Setelah menetapkan topik dan judul, langkah selanjutnya adalah identifikasi masalah, tahap pertama dalam proses penelitian yang memberikan landasan yang kokoh untuk penyelidikan lebih lanjut. Setelah merumuskan masalah, langkah berikutnya adalah menetapkan tujuan penelitian. Penelitian ini bertujuan untuk menerapkan Algoritma FCM (*Fuzzy C-Means*) untuk mengidentifikasi karakteristik donatur yang potensial untuk program ziswaf, serta Algoritma DBSCAN (*Density-Based Spatial Clustering Of Applications With Noise*) sebagai representasi pendekatan berdasarkan kepadatan untuk menemukan pendekatan yang paling sesuai. Selanjutnya, penelitian ini juga bertujuan untuk menerapkan Model LRFM untuk menentukan *Customer Lifetime Value* (CLV) donatur yang potensial. Tahap berikutnya adalah pengumpulan data, yang mencakup pengumpulan data yang relevan dan diperlukan untuk penelitian ini, terutama terkait dengan proses segmentasi data donatur. Setelah data terkumpul, dilakukan proses pembersihan data untuk mendapatkan data yang bersih. Selanjutnya, data dibentuk menjadi empat variabel: length, recency, frequency, dan monetary, untuk menghasilkan data LRFM. Nilai dalam data LRFM dinormalisasi untuk menyetarakan skala nilai dan menghasilkan data LRFM yang ternormalisasi. Selanjutnya, setiap variabel LRFM diberi bobot menggunakan metode Analytical Hierarchy Process. *Customer Lifetime Value* (CLV) adalah indeks yang menunjukkan nilai donatur bagi lembaga. CLV digunakan untuk mengevaluasi prioritas segmen donatur. Pada tahap perhitungan CLV, data LRFM yang telah dinormalisasi dikalikan dengan bobot masing-masing variabel dan menghasilkan data LRFM bobot untuk dimasukkan ke tahap *clustering*. Nilai K untuk *clustering* ditentukan menggunakan metode *Elbow*, *Calinski-*

Harabasz Index (CH), *Silhouette Coefficient*, dan *Davies Bouldin Index (DB)*. Data LRFM bobot kemudian di-cluster, hasil *clustering* divisualisasikan dan dianalisis karakteristik masing-masing segmen.

III. HASIL DAN PEMBAHASAN

A. Hasil Penelitian

1. Bussiness Understanding

Dalam penelitian ini, penulis menganalisis segmentasi donatur untuk mendukung pengembangan ide-ide kampanye donasi kreatif yang bertujuan meningkatkan pendapatan donasi. Data donasi yang tersimpan digunakan untuk mengelompokkan donatur berdasarkan parameter *length* (lama waktu sejak donasi terakhir), *recency* (waktu terakhir melakukan donasi), *frequency* (frekuensi donasi), dan *monetary* (jumlah nominal yang didonasikan). Penelitian ini menggunakan aplikasi pemrograman Python Streamlit untuk memproses pengelompokan dan pelabelan data donatur secara otomatis, sehingga memungkinkan identifikasi karakteristik unik dari setiap donatur.

2. Data Understanding

Pada tahapan ini, dilakukan analisis kebutuhan data terkait implementasi CRM (Customer Relationship Management) dan data mining untuk mencapai tujuan organisasi. Data yang digunakan mencakup rentang waktu dari bulan Maret 2020 hingga 11 Juli 2023. Laznas PPPA Daarul Qur'an memiliki sekitar 15.972 donatur perorangan dan total donasi tercatat sebanyak 697.450 data donasi.

3. Data Preparation

Tahapan persiapan data melibatkan semua aktivitas yang diperlukan untuk membangun dataset akhir.

a. Data Cleaning

Tahap pembersihan data melibatkan proses menghapus nilai null dan menyaring transaksi yang berhasil pada data dari pppa.id dan sedekaonline.com menggunakan aplikasi KNIME. Setelah data dibersihkan, dilakukan penggabungan data menjadi satu file dengan ekstensi CSV. Dataset yang dihasilkan terdiri dari 669.654 rekaman, dengan 15.972 donatur dan 641.414 transaksi.

b. Transformasi LRFM

Tahapan ini mencakup data transaksi donatur yang telah melewati proses pembersihan dan seleksi. Data kemudian diubah untuk mempermudah proses data mining. Pada tahap ini, data ditransformasi ke dalam nilai LRFM donatur. Parameter *Length*, *Recency*, *Frequency*, dan *Monetary* transaksi donatur dihitung dengan bantuan pemrograman Python.

c. Normalisasi

Normalisasi dan segmentasi data transaksi merupakan langkah penting sebelum dilakukan *clustering*. Normalisasi dilakukan untuk memastikan bahwa nilai variabel memiliki skala yang sama, sehingga tidak ada satu variabel yang mendominasi analisis. Segmentasi dilakukan untuk mengelompokkan donatur berdasarkan pola perilaku mereka, yang diukur berdasarkan empat kriteria yaitu *Length* (lama hubungan), *Recency* (kebaruan transaksi), *Frequency* (frekuensi transaksi), dan *Monetary* (nilai transaksi). Setiap kriteria diberi skor (*Length_Score*, *Recency_Score*, *Frequency_Score*, *Monetary_Score*) dan digabungkan menjadi *LRFM_Score*. Skor *LRFM_Score* kemudian digunakan untuk mengelompokkan donatur ke dalam segmen-segmen yang berbeda. Dengan hasil sebagai berikut :

Length's: Skew: -0.1667589718441922, : SkewtestResult(statistic=-8.54922764352441, pvalue=1.2391490434489838e-17)

Recency's: Skew: -1.1636350297960636, : SkewtestResult(statistic=-48.58453286348448, pvalue=0.0)

Frequency's: Skew: 0.5182486211512457, : SkewtestResult(statistic=25.221093147755, pvalue=2.351621919485632e-140)

Monetary's: Skew: -1.2255540303991999, : SkewtestResult(statistic=-50.38908450573138, pvalue=0.0)

4. Modeling

Penelitian ini bertujuan untuk mengukur bobot variabel LRFM (Length, Recency, Frequency, Monetary) dalam segmentasi pelanggan menggunakan metode AHP (Analytical Hierarchy Process). Data diperoleh dari kuesioner yang diisi oleh tiga orang responden. Bobot variabel LRFM dihitung dengan merata-ratakan nilai preferensi responden dan kemudian dinormalisasi. Selanjutnya, metode AHP digunakan untuk memproses data ternormalisasi dan menghasilkan bobot akhir untuk setiap variabel. Nilai eigen maksimum, yang merupakan representasi tingkat konsistensi data, dihitung dengan cara merata-rata nilai vektor konsistensi. Nilai eigen maksimum yang diperoleh, sebesar 7.17473545, selanjutnya digunakan sebagai parameter dalam kalkulasi Indeks Konsistensi (CI).

$$CI = \frac{7.17473545 - 7}{7 - 1} = 0.029122575$$

Kemudian dicari *consistency ratio* yaitu membagi *CI* dengan *RI*. Sehingga nilai *CI/RI* yang diperoleh adalah :

$$\frac{0.029122575}{1.32} = 0.022062557$$

Penelitian ini menguji validitas bobot kriteria yang digunakan dalam menghitung CLV. Hasil kuesioner menunjukkan nilai *CI/RI* sebesar 0.022, yang lebih kecil dari 0.1. Angka ini menunjukkan konsistensi internal kuesioner dan memvalidasi bobot kriteria yang diperoleh. Bobot kriteria untuk variabel length, recency, frequency, dan monetary berturut-turut adalah 0.1797, 0.1416, 0.2816, dan 0.3970. Kriteria monetary memiliki bobot terbesar, yaitu 0.3970, menunjukkan pengaruhnya yang paling signifikan terhadap nilai CLV.

5. Perhitungan CLV

Nilai *Customer Lifetime Value* (CLV) dihitung dengan mengalikan bobot yang telah ditentukan dengan nilai LRFM (*Loyalty, Recency, Frequency, Monetary*) masing-masing pelanggan dan menjumlahkan hasilnya.

6. Clustering FCM

Sebelum menerapkan algoritma *Fuzzy C-Means* (FCM), paragraf ini menjelaskan langkah awal menentukan jumlah kluster optimal (K) menggunakan metode K-Means. Metode "*elbow*" atau "*sum of squared errors*" (SSE) digunakan untuk menentukan nilai K yang optimal. Algoritma ini mengimplementasikan metode *Elbow Method* untuk menentukan jumlah kluster optimal dalam model K-Means. Pertama, *dictionary* dan list kosong dibuat untuk menyimpan nilai SSE dan inertia. Kemudian, *loop for* digunakan untuk iterasi melalui berbagai jumlah kluster (1 hingga 9). Di setiap iterasi, model KMeans dibuat dan dilatih dengan data, dan nilai inertia disimpan. Nilai inertia ini kemudian ditambahkan ke *dictionary* dan list. Terakhir, jumlah kluster yang menghasilkan nilai inertia terendah dipilih sebagai jumlah kluster optimal.

7. Clustering DBSCAN

Algoritma *clustering* DBSCAN diawali dengan penentuan nilai K optimal untuk jumlah kluster dalam metode KMeans menggunakan metode "*elbow*" atau "*sum of squared errors*" (SSE). Selanjutnya, dijelaskan langkah-langkah pembuatan script algoritma DBSCAN.

8. *Evaluation*

Evaluasi model dilakukan untuk mengukur performa klasifikasi algoritma FCM dan DBSCAN. Empat metrik digunakan, yaitu SSE (*Sum of Squared Error*), *Calinski-Harabasz Index*, *Silhouette Score*, dan *Davies Bouldin Index*.

9. *FCM*

Metrik-metrik seperti SSE, *Calinski-Harabasz Index*, *Silhouette Score*, dan *Davies Bouldin Index* berperan sebagai alat ukur untuk menilai kualitas kluster yang dihasilkan dalam analisis kluster data.

a. *Sum of Squared Errors (SSE)*

Analisis kluster bertujuan untuk mengelompokkan data berdasarkan karakteristik yang serupa. Salah satu metrik yang digunakan untuk mengukur kualitas pengelompokan adalah *Sum of Squared Error (SSE)*. SSE dihitung dengan menjumlahkan kuadrat jarak setiap titik data dalam sebuah kluster terhadap pusat klusternya. Semakin kecil nilai SSE, semakin baik pengelompokan data. Nilai SSE 4688563936.0715 menunjukkan total kuadrat jarak antara semua titik data dan pusat klusternya dalam hasil klusterisasi. Nilai ini dapat digunakan untuk mengevaluasi kualitas pengelompokan data dan membandingkannya dengan hasil klusterisasi lain.

b. *Calinski-Harabasz Index (CHI)*

CHI adalah metrik eksternal yang mengukur efektivitas klusterisasi dengan cara membandingkan dispersi antar-kluster dengan dispersi dalam-kluster. Nilai CHI yang tinggi menunjukkan klusterisasi yang optimal, di mana kluster-klusternya memiliki perbedaan yang jelas. Dalam konteks ini, nilai CHI sebesar 0.3128562606910544 mengindikasikan bahwa kluster-kluster yang terbentuk tidak terpisahkan dengan baik, sehingga membingungkan untuk membedakan antar kluster.

c. *Silhouette Score (SS)*

Sum of Squares (SS) adalah metrik yang mengukur kualitas pengelompokan data. Nilai SS berkisar antara -1 dan 1, dengan nilai yang lebih tinggi menunjukkan pengelompokan yang lebih baik. Dalam konteks ini, nilai SS 0.08061440564597909 menunjukkan adanya tumpang tindih atau ketidakjelasan dalam penempatan titik data ke dalam kluster. Hal ini menunjukkan bahwa pengelompokan data mungkin tidak optimal dan perlu dioptimalkan lebih lanjut.

d. *Davies Bouldin Index (DBI)*

Kualitas pengelompokan dapat diukur menggunakan *Indeks Davies-Bouldin (DBI)*, yang menilai tingkat separasi dan kohesi antar kluster. Nilai DBI yang rendah menunjukkan pengelompokan yang lebih baik. Pada analisis ini, DBI 9.910899528742423 menunjukkan tingkat separasi dan kohesi kluster yang rendah, mengindikasikan kualitas pengelompokan yang suboptimal.

10. *DBSCAN*

Analisis terhadap kinerja algoritma DBSCAN dengan menerapkan metrik-metrik yang telah ditentukan memungkinkan pengukuran kualitas kluster yang dihasilkan pada kumpulan data tertentu. Penjelasan lebih lanjut terkait hasil evaluasi akan dipaparkan pada bagian selanjutnya.

a. *Sum of Squared Errors (SSE)*

Dalam konteks DBSCAN, *Sum of Squared Errors (SSE)* tidak ideal sebagai metrik evaluasi utama. Hal ini dikarenakan DBSCAN tidak dirancang untuk mengoptimalkan SSE. Algoritma ini berfokus pada identifikasi area dengan kepadatan titik data tinggi, bukan meminimalkan jarak antar titik dalam kluster. Oleh karena itu, nilai SSE yang dihasilkan mungkin tidak memberikan informasi yang relevan atau mudah diinterpretasikan dalam konteks DBSCAN.

b. *Calinski-Harabasz Index (CHI)*

CHI (*Calinski-Harabasz Index*) adalah suatu metrik eksternal yang digunakan untuk mengevaluasi kualitas klusterisasi dengan cara menghitung rasio antara dispersi antar kluster dan dispersi dalam kluster. Nilai CHI yang tinggi menunjukkan klusterisasi yang baik, dengan kluster-kluster yang terdefinisi dengan jelas dan terpisah secara signifikan. Dalam studi ini, nilai CHI sebesar 3.23 juta menunjukkan bahwa klusterisasi yang dilakukan telah menghasilkan kluster-kluster yang terdefinisi dengan baik dan terpisah secara signifikan.

c. *Silhouette Score (SS)*:

Kualitas klusterisasi dapat diukur dengan *Silhouette Score (SS)* dan nilai SS yang tinggi, seperti 0.836, menunjukkan bahwa klusterisasi memiliki kualitas yang baik. Hal ini ditandai dengan kohesi internal yang tinggi, di mana setiap titik data dalam kluster memiliki kesamaan yang tinggi dengan titik data lainnya dalam kluster yang sama. Selain itu, separasi antar kluster juga tinggi, di mana kluster-kluster berbeda memiliki perbedaan yang signifikan satu sama lain. Kualitas klusterisasi yang baik juga ditandai dengan bentuk dan ukuran kluster yang optimal, di mana kluster-kluster tidak terlalu kecil atau terlalu besar, dan tidak memiliki bentuk yang aneh.

d. *Davies Bouldin Index (DBI)*

DBI merupakan metrik yang digunakan untuk mengevaluasi kualitas klusterisasi. Nilai DBI yang rendah menunjukkan pemisahan antar kluster yang baik serta kekompakan kluster itu sendiri. Dalam kasus ini, nilai DBI 2.19 menunjukkan kualitas klusterisasi yang baik. Lebih lanjut, nilai CHI dan SS juga mendukung kesimpulan ini, menunjukkan kualitas klusterisasi yang tinggi. Namun, perlu diingat bahwa interpretasi nilai-nilai ini bergantung pada karakteristik data dan konteks analisis. Secara keseluruhan, klusterisasi yang dihasilkan oleh DBSCAN pada data Anda menunjukkan kualitas yang baik.

B. Pembahasan

Penelitian ini berfokus pada evaluasi keefektifan metode klusterisasi *Fuzzy C-Means (FCM)* dan *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)* dalam segmentasi data donatur untuk PPPA Daarul Qur'an. Evaluasi dilakukan menggunakan beberapa indeks validasi klusterisasi yaitu *Sum of Squared Errors (SSE)*, *Calinski-Harabasz Index (CHI)*, *Silhouette Coefficient (SS)*, dan *Davies-Bouldin Index (DBI)*. Hasil evaluasi menunjukkan bahwa metode FCM menghasilkan SSE sebesar 4.563.140.439,7347, CHI sebesar 0,3128562606910544, SS sebesar 0,08061440564597909, dan DBI sebesar 9,910899528742423. Di sisi lain, metode DBSCAN menghasilkan CHI sebesar 3.237.127,1389106703, SS sebesar 0,8479515063332151, dan DBI sebesar 2,1939426200975047, dengan SSE tidak relevan untuk metode ini. Berikut ini adalah analisis hasil dari FCM:

1. *Sum of Squared Errors (SSE)*

SSE mengukur seberapa baik data sesuai dengan klasternya; nilai yang lebih rendah menunjukkan klusterisasi yang lebih baik. Nilai SSE yang tinggi pada FCM menunjukkan bahwa data tidak dikelompokkan secara efektif.

2. *Calinski-Harabasz Index (CHI)*

CHI menilai rasio antara total dispersal antar-kluster dengan total dispersal intra-kluster. Nilai CHI yang rendah pada FCM menunjukkan bahwa kluster yang dihasilkan kurang terpisah dengan baik dan memiliki dispersi internal yang tinggi.

3. *Silhouette Coefficient (SS)*

SS mengukur seberapa mirip objek berada dalam kluster yang sama dibandingkan dengan objek di kluster lain. SS berkisar antara -1 dan 1, di mana nilai mendekati 1 menunjukkan klusterisasi yang baik. Nilai SS yang sangat rendah pada FCM menunjukkan bahwa klusterisasi yang dilakukan tidak menghasilkan kluster yang jelas dan terpisah.

4. *Davies-Bouldin Index (DBI)*

DBI mengukur rata-rata kesamaan antara masing-masing kluster dengan kluster yang paling mirip, di mana nilai yang lebih rendah menunjukkan klusterisasi yang lebih baik. Nilai DBI yang tinggi pada FCM mengindikasikan bahwa kluster yang dihasilkan tidak kompak dan memiliki banyak overlap.

Berikut ini adalah analisis hasil DBSCAN:

1. *Calinski-Harabasz Index (CHI)*

Nilai CHI yang sangat tinggi pada DBSCAN menunjukkan bahwa kluster yang dihasilkan sangat terpisah dengan baik dan memiliki dispersi internal yang rendah, menandakan klusterisasi yang lebih efektif.

2. *Silhouette Coefficient (SS)*

SS yang mendekati 1 pada DBSCAN menunjukkan bahwa data dikelompokkan ke dalam kluster yang terpisah dengan sangat baik, dengan objek-objek dalam kluster yang sama sangat mirip satu sama lain.

3. *Davies-Bouldin Index (DBI)*

Nilai DBI yang rendah pada DBSCAN mengindikasikan bahwa kluster yang dihasilkan adalah kompak dan memiliki sedikit overlap, menunjukkan bahwa metode ini mampu menghasilkan kluster yang lebih homogen dan terpisah dengan baik.

4. *Perbandingan dan Kesimpulan*

Hasil penelitian tentang pengklasteran dan segmentasi karakteristik donatur sedekah daring memberikan kontribusi penting bagi PPPA Daarul Qur'an dalam pengelolaan dan strategi pemasaran organisasi mereka. Dengan memahami karakteristik yang berbeda dari setiap segmen donatur, PPPA Daarul Qur'an dapat mengoptimalkan pengelolaan sumber daya dan mengembangkan strategi pemasaran yang lebih efektif. Segmentasi yang lebih baik memungkinkan mereka untuk menyesuaikan pendekatan komunikasi dan kampanye dengan cara yang lebih personal dan relevan terhadap setiap segmen, meningkatkan resonansi pesan dan respons kampanye. Lebih lanjut, dengan memanfaatkan data ini, organisasi dapat meningkatkan efektivitas kampanye donasi mereka dengan cara yang lebih strategis, meningkatkan retensi donatur, menjangkau lebih banyak donatur potensial, dan mengukur kinerja kampanye dengan lebih baik. Ini semua bertujuan untuk meningkatkan jumlah donasi yang diterima dan dampak positif yang dihasilkan dalam masyarakat.

IV. KESIMPULAN

Berdasarkan rumusan masalah, tujuan penelitian, dan hipotesis, disimpulkan sebagai yaitu kategori donatur hanya didasarkan pada nominal, dengan tiga kelompok: donatur gold (di atas 1 juta), silver (100.000-1 juta), dan bronze (di bawah 100.000). Namun, setelah analisis dataset, hasil kinerja *clustering Fuzzy C-Means* (FCM) dan DBSCAN menunjukkan *Clustering FCM*, berdasarkan nilai $K = 2$ dari metode Elbow yang optimal, menggunakan inisialisasi matriks bobot acak dari dataset hasil LRFM, menghasilkan dua kluster. *Clustering DBSCAN* juga menghasilkan dua kluster dan membentuk lima rentang penghasilan tahunan: 1000-2000, 2000-3000, 3000-4000, 4000-5000, dan 5000-6000. Kluster yang paling sesuai dengan profil donatur potensial dipilih berdasarkan hasil ini. Hasil evaluasi menunjukkan bahwa algoritma DBSCAN memberikan klusterisasi yang lebih baik dibandingkan dengan *Fuzzy C-Means* (FCM) pada dataset penelitian ini, dengan pemisahan dan homogenitas kluster yang lebih superior. Penelitian ini juga menunjukkan bahwa pemodelan LRFM dapat menentukan *Customer Lifetime Value* (CLV) dari donatur yang potensial.

Pemahaman dan perhitungan CLV membantu lembaga PPPA Daarul Quran dalam mengidentifikasi donatur bernilai tinggi, merencanakan strategi pemasaran yang lebih efektif, dan mengoptimalkan pengeluaran untuk akuisisi donatur baru. Berdasarkan data LRFM dan CLV, terbentuk enam segmen dari delapan segmen yang diidentifikasi, yaitu: Donatur Loyal,

Donatur Potensial, Donatur Menjanjikan, Donatur Butuh Perhatian, Donatur Tertidur, dan Donatur Berisiko Hilang. Strategi komunikasi dapat dirancang sesuai dengan karakteristik klaster dan segmen yang telah diidentifikasi, menggunakan bahasa dan pendekatan yang sesuai dengan minat para donatur. Tim divisi CRM dapat membuat kampanye atau kegiatan khusus yang ditargetkan kepada setiap klaster, dengan menyesuaikan pesan, promosi, atau tawaran program ZISWAF berdasarkan karakteristik klaster. Selain itu, penting untuk membangun hubungan jangka panjang dengan donatur potensial, bukan hanya untuk satu kali donasi. Terakhir, pantau respons dan partisipasi dari setiap klaster dan lakukan evaluasi secara berkala untuk mengukur efektivitas strategi serta mengidentifikasi perubahan yang diperlukan.

Penelitian ini memiliki beberapa kekurangan yang perlu dikembangkan lebih lanjut. Pengembangan yang disarankan meliputi penambahan model algoritma klustering lainnya dan membandingkan metode data mining yang berbeda. Selain itu, penggunaan perangkat lunak atau alat lain yang dapat mendukung proses klustering data juga disarankan. Penting untuk memastikan bahwa penggunaan data untuk tujuan pemasaran atau pendekatan pribadi mematuhi kebijakan privasi dan hukum yang berlaku. Evaluasi dan penyempurnaan strategi perlu dilakukan secara berkelanjutan berdasarkan respons dan hasil yang diperoleh.

REFERENSI

- [1] F. Hariyadi, H. Sujaini dan A. Sukamto, "Pengklasteran Mahasiswa UNTAN Berdasarkan IPK Menggunakan Metode Fuzzy C-Means," *JUSTIN: Jurnal Sistem dan Teknologi Informasi*, vol. 1, no. 1, pp. 170-180, 2020.
- [2] A. Supriadi dan S. Moh, "Clustering Loyalitas Pelanggan Dengan Model RFM (Recency, Frequency, Monetary) Dan Metode K-Means Berbasis Median," *UNUJA*, vol. 2, no. 2, pp. 1-20, 2021.
- [3] B. N. Sar dan A. Primajaya, "Penerapan Clustering DbSCAN Untuk Pertanian Padi Di Kabupaten Karawang," *Jurnal Informatika dan Komputer*, vol. 1, no. 1, pp. 1-20, 2019.
- [4] Sudriyanto, "Clustering loyalitas pelanggan dengan metode rfm (recency, frequency, monetary) dan fuzzy c-means," *Prosiding SNATIF*, vol. 1, no. 2, pp. 815-822, 2019.
- [5] D. B. Saputra dan E. Riksakomara, "Implementasi Fuzzy C-Means dan Model RFM untuk Segmentasi Pelanggan (Studi Kasus: PT. XYZ)," *Jurnal Teknik ITS*, vol. 7, no. 1, pp. 119-124, 2019.
- [6] Alfian, "Clustering Kesetiaan Pelanggan E-Ritel Dengan Model RFM (Recency, Frequency, Monetary) dan K-Means," *Jurnal Teknologi Informasi*, vol. 14, no. 2, pp. 239-250, 2019.
- [7] N. Syarif dan W. Windarto, "Implementasi Algoritma Fuzzy C-Means Dan Metode Recency Frequency Monetary (Rfm) Pada Aplikasi Data Mining Untuk Pengelompokan Pelanggan," *Sebatik*, vol. 22, no. 2, pp. 88-94, 2019.
- [8] A. Syafrianto, "Perancangan Aplikasi K-Means Untuk Pengelompokan Mahasiswa STMIK ELRAHMA Yogyakarta Berdasarkan Frekuensi Kujungan KePerpustakaan dan IPK," *Jurnal Informatika*, vol. 1, no. 1, pp. 1-20, 2019.
- [9] B. Nurina dan A. Primajaya, "Penerapan Clustering DbSCAN Untuk Pertanian Padi Di Kabupaten Karawang," *Jurnal Informatika dan Komputer*, vol. 4, no. 1, pp. 1-20, 2019.
- [10] M. Azlansahputra, "Penerapan Algoritma Fuzzy C-Means Menggunakan Model RFM Dalam Klasterisasi Pelanggan Pada Toko Kue Feandra Cake," *Jurnal Ilmiah Rekayasadan Manajemen Sistem Informasi*, vol. 6, no. 1, p. 64-69, 2020.
- [11] M. Gupta dan P. Chandra, "A comprehensive survey of data mining," *International Journal of Information Technology*, vol. 12, no. 4, pp. 1243-1257, 2020.

- [12] A. Dogan dan D. Birant, "Machine learning and data mining in manufacturing," *Expert Systems with Applications*, vol. 1, no. 10, pp. 345-350, 2021.
- [13] J. Kong, C. Yang, J. Wang, X. Wang dan M. Zuo, "Deep-stacking network approach by multisource data mining for hazardous risk identification in IoT-based intelligent food management systems," *Computational Intelligence and Neuroscience*, vol. 1, no. 21, pp. 119-130, 2021.
- [14] S. Monalisa dan R. Zafa, "Klasterisasi Customer Lifetime Value Dengan Model LRFM Menggunakan Algoritma Fuzzy C-Means," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 6, no. 1, pp. 38-42, 2020.
- [15] L. Waroka, S. Monalisa, D. Anjainah dan N. Arifin, "Implementasi Algoritma Fuzzy C-Means (Fcm) Dalam Pengklasterisasian Nilai Hidup Pelanggan Dengan Model Lrfm.," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 6, no. 1, pp. 1-20, 2020.
- [16] M. Migdadi, "Knowledge management, customer relationship management and innovation capabilities," *Journal of Business & Industrial Marketing*, vol. 36, no. 1, pp. 111-124, 2020.
- [17] C. Zhang, X. Wang, A. Cui dan S. Han, "Linking big data analytical intelligence to customer relationship management performance," *Industrial Marketing Management*, vol. 9, no. 1, pp. 483-494, 2020.
- [18] P. Bhattacharjee dan P. Mitra, "A survey of density based clustering algorithms," *Frontiers of Computer Science*, vol. 15, no. 1, pp. 1-27, 2021.
- [19] S. Zhou, D. Li, Z. Zhang dan R. Ping, "A new membership scaling fuzzy c-means clustering algorithm," *IEEE Transactions on Fuzzy Systems*, vol. 29(9), no. 9, pp. 2810-2818, 2020.
- [20] Y. Tang, F. Ren dan W. Pedrycz, "Fuzzy C-means clustering through SSIM and patch for image segmentation," *Applied Soft Computing*, vol. 8, no. 7, pp. 230-245, 2020.
- [21] G. Oyewole dan G. Thopil, "Data clustering: application and trends," *Artificial Intelligence Review*, vol. 56, no. 7, pp. 6439-6475, 2023.
- [22] W. Kurniawan dan S. Rahayu, "Analisis Segmentasi Usia Donatur Terhadap Peminatan Donasi Di Human Initiative," *Manajemen Dewantara*, vol. 7, no. 3, pp. 457-464, 2023.
- [23] T. Saranya dan S. Saravanan, "Groundwater potential zone mapping using analytical hierarchy process (AHP) and GIS for Kancheepuram District, Tamilnadu, India," *Modeling Earth Systems and Environment*, vol. 6, no. 2, pp. 1105-1122, 2020.