



Pemetaan Pelanggan dengan LRFM dan Two Stage Clustering untuk Memenuhi Strategi Pengelolaan

Ni Putu Viona Viandari¹, I Made Agus Dwi Suarjaya², I Nyoman Piarsa³

^{1,2,3} Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Udayana

¹ vionaviandarii19@gmail.com, ² agussuarjaya@it.unud.ac.id, ³ manpits@unud.ac.id

Abstract

Maibus is a company of transportation services located in Bali. Transaction data that is owned has not been managed properly. This results in data accumulation and only as a turnover calculation, so LRFM and clustering methods are needed to assist the calculation and processing data in fulfilling customer management strategies. The research was conducted by collecting and understanding data, preprocessing, applying LRFM (Length, Recency, Frequency, Monetary), normalizing LRFM, evaluating the number of clusters with Davies Bouldin Index (DBI), clustering with K-Means, and analyzing cluster results. The data used is transaction data from January 2017 to December 2018 with a total of 14.292 data. The clustering method with the K-means algorithm helps in mapping customers based on transaction data. DBI was used to determine the optimal number of clusters and LRFM used to test the determination of variables in determining customer behavior and loyalty. The results of testing 7.193 invoice using 5 clusters with DBI value is 0.135. The result of customers in cluster 0,2,4 are new customer groups with the proposed strategy is enforced strategy, while the customer in cluster 1 and 3 are lost customers with the proposed strategy is let-go strategy that refers to the customer value and customer loyalty matrix.

Keywords: LRFM, Clustering, Davies Bouldin Index, Customer Mapping, Data Mining

Abstrak

Maibus merupakan perusahaan di bidang jasa transportasi yang berlokasi di pulau Bali. Data transaksi yang dimiliki perusahaan belum dikelola dengan baik. Hal itu mengakibatkan penumpukan data dan hanya sebagai perhitungan omset, maka diperlukan metode LRFM (*Length, Recency, Frequency, Monetary*) dan *clustering* untuk membantu perhitungan dan pengolahan data pelanggan dalam memenuhi strategi pengelolaan pelanggan. Penelitian dilakukan dengan mengumpulkan dan memahami data, *preprocessing*, menerapkan LRFM, normalisasi LRFM, evaluasi jumlah *cluster* dengan *Davies Bouldin Index*, *clustering* dengan K-Means, dan analisis hasil *cluster*. Data yang digunakan yaitu data transaksi dua periode Januari 2017 hingga Desember 2018 dengan total 14.292 data. Metode *clustering* dengan algoritma K-means membantu dalam pemetaan pelanggan berdasarkan data transaksi. *Davies Bouldin Index* dilakukan untuk menentukan jumlah *cluster* yang optimal dan LRFM digunakan sebagai pengujian penentuan variabel dalam menentukan perilaku dan loyalitas pelanggan. Tujuan penelitian ini dilakukan untuk mengetahui kelompok dan jenis pelanggan yang memiliki potensial menguntungkan bagi perusahaan dan mengusulkan strategi marketing yang tepat bagi masing – masing kelompok pelanggan. Pengujian 7.193 data faktur transaksi dengan menggunakan hasil penentuan jumlah *k-cluster* yang optimal yaitu 5 *cluster* dengan nilai *Davies Bouldin Index* adalah 0.135 menghasilkan dua kelompok pelanggan dan dua strategi *marketing* yang berdasarkan *customer value matrix* dan *customer loyalty matrix*. Hasil yang diperoleh dalam pengelompokan pelanggan pada *cluster 0, cluster 2, cluster 4* merupakan kelompok pelanggan *new customer* dengan strategi yang diusulkan adalah *enforced* strategi, sedangkan kelompok pelanggan pada *cluster 1, cluster 3* termasuk kelompok pelanggan *lost customer* dengan strategi yang diusulkan adalah *let-go* strategi.

Kata kunci: LRFM, *Clustering*, *Davies Bouldin Index*, Pemetaan Pelanggan, *Data Mining*

1. Pendahuluan

Bali merupakan tempat tujuan wisata yang banyak pengunjungnya. Tidak heran jika banyak jutaan wisatawan asing maupun domestik berkunjung ke pulau Bali dikarenakan keindahan alam, adat dan budayanya.

Dengan data yang diperoleh menyatakan bahwa pada tahun 2019 wisatawan yang berkunjung ke pulau Bali meningkat dibandingkan dengan tahun sebelumnya yaitu tahun 2018 [1]. Semakin banyaknya kunjungan wisatawan ke pulau Bali, hal tersebut berdampak pula bagi perusahaan yang bergerak dibidang penyewaan

bus. Perusahaan PT. Nadia Kencana merupakan salah satu perusahaan yang bergerak di bidang jasa transportasi pariwisata di Bali [2]. Perusahaan tersebut saat ini telah memiliki banyak *customer* yang memilih untuk menyewa bus pada perusahaan tersebut. Dalam mempertahankan keberadaannya dan menghadapi pesaingnya di era pemasaran modern saat ini, perusahaan berupaya untuk fokus dalam pengelolaan pelanggan. *Customer Relationship Management* (CRM) yang dimana merupakan salah satu strategi bisnis dengan mengelola informasi rinci dari masing – masing pelanggan dan secara cermat mengelola hubungan terhadap pelanggan demi memaksimalkan kesetiaan pelanggan [3]. PT. Nadia Kencana memiliki permasalahan terkait dengan manajemen hubungan pelanggan dimana belum adanya informasi dalam menentukan pelanggan yang dinilai memiliki potensial menguntungkan dan pelanggan yang dinilai kurang memiliki potensial menguntungkan bagi perusahaan. Perusahaan hanya melakukan pendataan dari data transaksi melalui data pelanggan yang sudah terdaftar. Hal tersebut mengakibatkan menumpuknya data dan hanya dijadikan sebagai perhitungan omzet bagi perusahaan. Salah satu cara yang dapat dilakukan yaitu dengan teknik data mining atau yang merupakan suatu proses untuk menemukan informasi dari beberapa kumpulan data [4] yang bertujuan untuk membuat keputusan bisnis yang lebih efisien dan mengambil keputusan secara tepat [5]. Teknik data mining dengan metode *clustering* merupakan metode pengelompokan record, pengamatan, dan membentuk kelas objek – objek yang memiliki kemiripan, dimana metode *clustering* digunakan oleh penulis dalam penelitian ini [6]. Penelitian ini menggunakan data transaksi berupa excel yang diberikan langsung oleh pihak perusahaan, dimana pada data transaksi tersebut terdapat field data yang dapat digunakan untuk melakukan pemetaan pelanggan sehingga diperlukan metode algoritma yang tepat dalam pengelompokan data.

Penelitian yang berhubungan dengan pemetaan pelanggan potensial yaitu oleh Yulison Herry Chrisnanto, dimana proses teknik data mining dengan metode *clustering* K-Means dapat digunakan dengan baik untuk melakukan segmentasi loyalitas pelanggan menggunakan model analisis RFM (*Recency, Frequency, Monetary*). Hasil pengelompokan dilakukan dengan 4 *cluster* dan melalui model RFM tersebut dapat disimpulkan bahwa 215 pelanggan dengan tingkat loyalitas cukup baik dan 138 pelanggan dengan tingkat loyalitas rendah [7].

Penelitian selanjutnya adalah Analisis pemetaan pelanggan potensial menggunakan algoritma K-Means dan LRFM model untuk mendukung strategi pengelolaan pelanggan menjadi salah satu kategori dalam topik judul yang berdasarkan isi dari materi.

Penelitian ini menghasilkan kinerja algoritma K-Means dan LRFM sangat baik dalam menganalisis data pelanggan dan diperoleh hasil jumlah *cluster* yaitu 4 dan memiliki karakteristik pelanggan yang berbeda, dimana hasil tersebut bisa memberikan strategi *marketing* bagi perusahaan kepada 4 segmen pelanggan tersebut [8].

Penelitian yang terkait selanjutnya oleh Siti Monalisa yang dimana pada penelitian ini CLV (*Customer Lifetime Value*) digunakan dalam menentukan loyalitas pelanggan dengan model LRFM dan algoritma K-Means. Penelitian tersebut menghasilkan 3 *cluster* terbaik dengan metode validasi *cluster dunn index* dan *Silhouette Coefficient*. Berdasarkan jumlah 3 *cluster* tersebut menghasilkan nilai CLV yang tertinggi pada *cluster* 2 dengan nilai 0,208 sehingga segmen pelanggan pada *cluster* 2 memiliki loyalitas yang tinggi dan perlu dipertahankan dengan memberikan pelayanan terbaik [9].

Penelitian selanjutnya yaitu oleh Silvi Eka Susanty, dimana pada penelitian ini menggunakan proses data mining dalam melakukan segmentasi pelanggan untuk mendukung strategi pengelolaan pelanggan. Penelitian ini menggunakan model LRFM dan two stage clustering. Tahapan *two stage clustering* adalah tahapan yang dimulai dengan mencari jumlah *cluster* yang optimal dengan metode ward's yang digunakan sebagai inputan k-means dalam membentuk k-*cluster* dan tahapan kedua yaitu proses *clustering* dengan algoritma K-Means. Hasil yang diperoleh dari penelitian ini adalah terbentuknya 4 segmen pelanggan dari proses pemetaan LRFM *index* ke dalam irisan *customer value matrix* dan *customer loyalty matrix*. Keempat segmen pelanggan tersebut terdiri dari *Core Customers* dengan jenis *Platinum Customers* pada *cluster* 1, *Lost Customers* dengan jenis *Uncertain Lost Customers* pada *cluster* 2, *New Customers* dengan jenis *Uncertain New Customers* pada *cluster* 3 dan *Core Customers* dengan jenis *High Frequency Buying Customers* pada *cluster* 4. Pelanggan yang berperan besar dalam memberikan kontribusi pada perusahaan yaitu pada *core customer* [4].

Penelitian berikutnya yang masih terkait dengan segmentasi pemetaan pelanggan yaitu oleh Fitri Marisa, dkk, dimana pada penelitian ini menggunakan clustering K-Means dan model LRFM. Dalam mencari jumlah *cluster* terbaik pada penelitian ini dilakukan dengan metode elbow dan Sum Square Error (SSE) dan menghasilkan jumlah k-*cluster* adalah 2. Setelah melalui tahapan *clustering* dengan K-Means kemudian hasil tersebut dilakukan analisis dengan CLV untuk mengetahui nilai profitabilitas dan loyalitas pelanggan dengan model LRFM. Hasil akhir yang diperoleh yaitu dengan nilai CLV tertinggi berada di *cluster* 2, dimana nilai LRFM yang dihasilkan yaitu *length* yang tinggi, *frequency* yang tinggi, *recency* yang rendah dan

monetary yang lebih tinggi atau dengan kata lain bahwa pelanggan pada segmen ini sering melakukan transaksi dan jumlah uang yang dikeluarkan cukup tinggi, maka dapat dikatakan pelanggan ini memiliki profitabilitas yang tinggi dan perlu dipertahankan agar tidak menjadi incaran bagi para pesaing [10].

Berdasarkan beberapa penelitian yang telah dijabarkan sebelumnya dapat ditarik kesimpulan bahwa metode *clustering* dengan K-Means dinilai bisa memberikan hasil yang akurat dan optimal dalam mengelompokkan data dengan karakteristik yang sama dan berbeda dari kelompok data yang lain, sedangkan model RFM (*Recency, Frequency, Monetary*) yang dimana terdapat penambahan modifikasi pada variable *Length* atau disebut dengan model LRFM oleh Chang dan Tsay (2004) yaitu untuk menilai lamanya hubungan antara pelanggan dengan suatu perusahaan.

Metode *clustering* dan model LRFM merupakan perpaduan dua metode yang memiliki hubungan erat dalam mendapatkan hasil kelompok pelanggan berdasarkan karakteristik dari nilai atribut LRFM. Dalam proses *clustering*, model LRFM sangat diperlukan untuk mengetahui nilai loyalitas para pelanggan yang dilihat dari nilai atribut LRFM seperti melihat waktu terakhir transaksi yang dilakukan oleh pelanggan (*recency*), seberapa sering pelanggan tersebut melakukan transaksi (*frequency*), seberapa banyak nilai jumlah dalam satuan rupiah yang dilakukan oleh pelanggan saat melakukan transaksi (*monetary*), dan seberapa lama hubungan yang terjalin antara pelanggan dengan perusahaan (*length*) yang dihasilkan pada masing – masing pelanggan. Oleh karena itu jika proses *clustering* dilakukan tanpa menggunakan model LRFM maka perolehan hasil dalam menentukan kelompok pelanggan yang memiliki loyalitas dan berpotensi menguntungkan bagi perusahaan tidak akan berhasil.

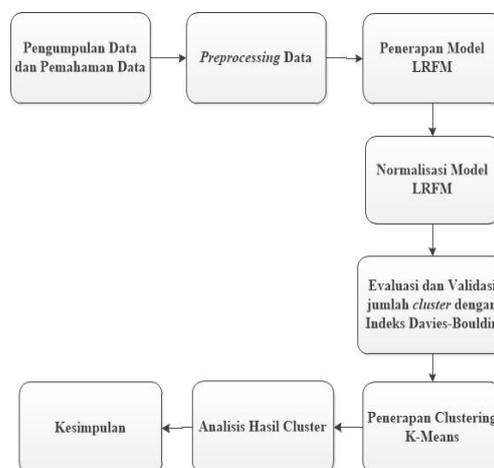
Pada penelitian ini penulis menggunakan dua metode yaitu metode *clustering* dengan algoritma K-Means dan model LRFM. Metode *clustering* dipilih oleh penulis karena dapat membantu dalam menentukan kelompok yang tidak diketahui dalam suatu data sedangkan model LRFM dipilih untuk digunakan dalam mengetahui nilai loyalitas dari masing – masing pelanggan dalam menentukan pelanggan yang bernilai potensial dan tidak berpotensi bagi perusahaan. Model LRFM dipilih dibandingkan dengan model RFM karena adanya penambahan atribut *length* yang dapat mengetahui seberapa lama hubungan yang telah berjalan antara pelanggan dengan perusahaan. Dalam pengoptimalan jumlah *k-cluster* yang dilakukan oleh penulis pada penelitian ini yaitu menggunakan validasi Davies Bouldin *Index*, dimana proses validasi ini memiliki perbedaan pada penelitian sebelumnya. Data transaksi yang digunakan oleh penulis pada penelitian ini merupakan data transaksi yang diberikan oleh

perusahaan PT. Nadia Kencana, dimana data transaksi ini berbeda dengan data transaksi yang digunakan oleh peneliti sebelumnya yang telah melakukan penelitian. Proses dalam melakukan *preprocessing* data yaitu dengan menggunakan bahasa pemrograman Python dan proses metode *clustering* algoritma K-Means dilakukan dengan bantuan *tools* yaitu RapidMiner oleh penulis. Dalam perolehan nilai model LRFM penulis menggunakan bantuan *query* pada SQL. Perolehan hasil dari pengelompokan pelanggan melalui proses *clustering* dan model LRFM kemudian disesuaikan dengan *customer value* dan *customer loyalty matrix* untuk mendapatkan kelompok dan jenis pelanggan dari setiap *cluster* yang telah terbentuk dan selanjutnya dapat memberikan usulan strategi marketing yang sesuai dengan kelompok dan jenis pelanggan dengan mengacu kepada nilai atribut *frequency* dan *monetary*.

Tujuan penulis pada penelitian ini dengan menggunakan metode *clustering* dan model LRFM adalah untuk dapat mengelompokkan para pelanggan di kelompok yang memiliki karakteristik yang sama dan berbeda dari kelompok pelanggan lainnya, sehingga dapat mengetahui masing – masing kelompok pelanggan yang berpotensi menguntungkan dan yang tidak banyak berkontribusi bagi perusahaan, dimana hasil dari masing – masing kelompok pelanggan yang sudah terbentuk dapat memberikan usulan *marketing* bagi perusahaan berdasarkan untuk bisa mengelola dan menjaga hubungannya kepada para pelanggan.

2. Metode Penelitian

Metode penelitian yang dilakukan yaitu dengan melalui enam tahapan untuk dapat mendapatkan hasil. Dapat dilihat pada Gambar 1, proses yang dilakukan dari awal yaitu pengumpulan dan pemahaman data, *preprocessing* data, penerapan model LRFM, normalisasi nilai LRFM, evaluasi dan validasi jumlah *cluster* dengan Indeks Davies Bouldin, penerapan *clustering* K-Means, dan terakhir adalah menganalisis hasil *cluster* [8].



Gambar 1. Metode Penelitian

2.1. Pengumpulan dan Pemahaman Data

Data yang digunakan pada penelitian dikumpulkan dari perusahaan PT. Nadia Kencana. Data yang diperoleh merupakan data transaksi periode Januari 2017 hingga Desember 2018. Rumus ditulis secara jelas menggunakan *equation editor* dengan nomor indeks seperti rumus 1.

Data tersebut merupakan rekapan data excel sebanyak dua folder yang dimana masing-masing folder terdapat dua belas dokumen yang berisi rekapan transaksi harian dalam satu bulan. Pemahaman data dari terkumpulnya data tersebut dilakukan untuk dapat memilah field data yang bisa digunakan pada proses selanjutnya.

2.2. Preprocessing Data

Dilakukannya *preprocessing* data yaitu dengan tujuan untuk mengubah data mentah ke dalam suatu format yang prosesnya lebih mudah dan efektif [11].

Integrasi data yang dilakukan pada penelitian ini yaitu data transaksi yang telah disaring dan disatukan untuk mendapatkan poin – point penting dengan berupa file CSV yang telah dikonversikan dari file excel sebelumnya, kemudian digabungkan ke dalam satu basis data pada SQL Server untuk dapat digunakan pada proses selanjutnya.

Reduksi data yang dilakukan pada penelitian ini yaitu dengan menghilangkan kolom atribut pada data yang tidak diperlukan sehingga dapat menghasilkan data input untuk proses *clustering* [11]. Pada penelitian ini terdapat atribut data yang dihilangkan dan atribut yang digunakan serta catatan yang dihilangkan karena terdapat *missing value*. *Missing value* merupakan informasi yang tidak tersedia untuk sebuah objek atau kasus. *Missing value* dapat terjadi karena informasi dari objek sulit untuk dicari, tidak diberikan, atau memang informasi tersebut tidak ada.

Missing value pada dasarnya tidak bermasalah bagi keseluruhan data, jika jumlahnya hanya sedikit, semisal hanya 1 % dari keseluruhan data. Namun jika persentase data yang hilang cukup besar, maka perlu dilakukan pengujian apakah data yang memiliki banyak *missing* tersebut masih layak diproses lebih lanjut atau tidak [12].

Setelah mendapatkan atribut data yang dapat digunakan dari proses reduksi selanjutnya adalah pembersihan data atau data *cleaning* yang memiliki tujuan untuk menghilangkan nilai - nilai yang tidak diperlukan, memperbaiki kekacauan data serta memeriksa data yang dianggap tidak konsisten [13]. Pada penelitian ini pembersihan data dilakukan dengan mengambil salah satu atribut yaitu nama pelanggan, dimana di dalam nama pelanggan tersebut terdapat kata yang memiliki arti sama dan beberapa karakter yang tidak diperlukan.

Dalam hal ini data transaksi yang diberikan tidaklah lengkap karena dalam proses selanjutnya diperlukan atribut lain seperti ID *Customer*, maka dari itu perlunya tabel pelanggan dengan atribut ID *Customer* pada penelitian ini. Pembuatan tabel pelanggan dengan menggunakan bahasa pemrograman Python ini terkoneksi langsung ke dalam basis data yang telah dibuat pada SQL Server.

2.3. Penerapan Model LRFM

Hasil data yang telah diperoleh pada *preprocessing* data merupakan data yang siap untuk diolah. Pada penerapan LRFM ini, data transaksi yang berhubungan dengan masing – masing atribut LRFM akan ditransformasikan ke dalam masing – masing atribut LRFM [8,14,15]. Transformasi data ke dalam atribut LRFM dapat dilihat pada Tabel 1.

Tabel 1. Transformasi Atribut LRFM

Atribut <i>Frequency</i>	Atribut <i>Monetary</i>	Atribut <i>Recency</i>	Atribut <i>Length</i>
Jumlah data transaksi yang dilakukan selama periode berlangsung	Jumlah nilai transaksi selama periode berlangsung	Tanggal terakhir dari transaksi	Jumlah waktu dalam satuan hari, sejak transaksi pertama hingga transaksi terakhir

2.4. Normalisasi LRFM

Normalisasi dari hasil masing – masing nilai pada atribut LRFM dilakukan karena untuk menghindari permasalahan mengenai rentang data yang terlalu luas dan agar dapat lebih dipahami dan mempermudah dalam proses *clustering*. Normalisasi dilakukan dengan metode *Z-score*. *Z-score* merupakan salah satu metode normalisasi yang berdasarkan nilai rata – rata (*mean*) dan standar deviasi dari suatu data [16]. Normalisasi dengan *z-score* dipilih oleh penulis pada penelitian ini karena dapat membantu dalam penentuan sebuah data apakah terdapat data *outlier*, dimana data *outlier* merupakan data dengan rentang nilai rata – ratanya sangat jauh. Rancangan dari *z-score* digunakan untuk mengolah skor dari variasi data set dan menjadikan skor tersebut sebagai proses dalam mengkomperasikan data lainnya secara akurat. Normalisasi dengan *Z-score* dihitung dengan Persamaan (1)

$$v' = \left(\frac{v_i - E_i}{std(E)} \right) \quad (1)$$

dengan atribut v' dikatakan sebagai hasil dari nilai normalisasi, kemudian v_i merupakan nilai yang akan dinormalisasi dalam atribut. Atribut E_i merupakan nilai rata – rata dari atribut dan $std(E)$ merupakan atribut standar deviasi E [17].

2.5. Evaluasi jumlah Cluster dengan DBI

Dalam proses *clustering*, diperlukannya pengoptimalan dalam jumlah *cluster* yang diinginkan agar mendapatkan hasil yang terbaik dan optimal. Diumpamakan jika $Sc(k)$ merupakan jarak intra-kelompok dalam kelompok k . Perumusan menghitung $Sc(k)$ dapat dilihat pada Persamaan (2)

$$Sc(k) = \frac{\sum_i^{N_k} |X_i - C_k|^2}{N_k} \quad (2)$$

dengan N_k adalah banyak titik dalam kelompok k , C_k merupakan *centroid* dari kelompok k , dan X_i merupakan data ke- i dalam kelompok k .

Berdasarkan dari tipe penelitiannya tahapan kegiatan perlu dijabarkan. Bagaimana penelitian tersebut dilakukan dan data yang akan diperoleh, perlu dijabarkan dalam bagian ini. Perumusan dalam pendefinisian jarak inter-kelompok dijabarkan pada Persamaan (3)

$$d_{kl} = |C_k - C_l| \quad (3)$$

dengan d_{kl} merupakan jarak antara kelompok k dan kelompok l , C_k adalah *centroid* kelompok k , dan C_l adalah *centroid* kelompok l [18]. Dapat diketahui bahwa perhitungan nilai Indeks Davies Bouldin secara jelas seperti pada Persamaan (4)

$$DB(nc) = \frac{1}{nc} \sum_{k=1}^{nc} \max \left(\frac{sc(k) + sc(l)}{d_{kl}} \right) \quad (4)$$

2.6. Penerapan Clustering K-Means

K-Means merupakan metode yang memisahkan data ke dalam *cluster*, sehingga data dengan karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data dengan karakteristik yang berbeda dikelompokkan ke dalam *cluster* lain [19]. Tahapan dari metode algoritma K-Mean dapat dilihat sebagai berikut:

1. Pilih objek k secara acak sebagai dasar pusat *cluster* (*centroid* awal) sebanyak k
2. Tentukan objek data selain *centroid* berdasarkan kedekatan setiap objek dengan *centroid*nya
3. Tentukan *centroid* baru dari *cluster* melalui nilai rata-rata setiap nilai objek pada setiap *cluster*
4. Ulangi langkah ke-2 dan ke-3 hingga nilai *centroid* tidak berubah [7].

Pencarian dalam menghitung jarak titik data objek ke titik data pusat (*centroid*) yaitu dengan menggunakan rumus *Euclidean Distance*. Perumusan dalam menghitung *Euclidean Distance* dapat dilihat pada Persamaan (5)

$$d_{Euclidean}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (5)$$

dengan $d(x,y)$ merupakan jarak antara data pada titik x dan y , x adalah titik data objek, y adalah titik data *centroid*, dan i adalah jumlah atribut data [20].

Penerapan metode *clustering* dengan algoritma K-Means pada penelitian ini, penulis menggunakan bantuan *tools* RapidMiner. Dengan *tools* pada RapidMiner, urutan dari tahapan algoritma K-Means akan diproses dengan parameter, operator dan fungsi yang mendukung pada metode *clustering* data.

2.7. Analisis Hasil Cluster

Hasil dari perolehan *clustering* selanjutnya dilakukan analisa untuk dapat mengetahui jenis dan kelompok pelanggan pada masing – masing *cluster* yang terbentuk dengan melihat dari indeks LRFM.

Perhitungan standar deviasi dilakukan untuk mendapatkan nilai indeks dari masing – masing atribut LRFM. Perolehan yang dihasilkan dari indeks LRFM tersebut kemudian dilakukan pemetaan pelanggan berdasarkan *customer loyalty matrix* dan *customer value matrix* [4]. Grup pelanggan menurut kelompok dan jenisnya dapat dilihat pada Tabel 2.

Tabel 2. Grup Pelanggan

No	Kelompok Pelanggan	Jenis Pelanggan
1	Core Customer (CC)	Terdiri dari : 1. <i>Including high value loyal customers</i> (LRFM, ↑↑↑)
		2. <i>High Frequency buying customers</i> (LRFM, ↑↑↓) 3. <i>Platinum customers</i> (LRFM, ↑↓↑)
2	Potential Customers (PC)	Terdiri dari : 1. <i>Potential loyal customers</i> (LRFM, ↑↑↑)
		2. <i>Potential high frequency customers</i> (LRFM, ↑↑↓) 3. <i>Potential consumption customers</i> (LRFM, ↑↓↑)
3	Lost Customers (LC)	Terdiri dari : 1. <i>High value lost customers</i> (LRFM, ↓↑↑)
		2. <i>Frequency lost customers</i> (LRFM, ↓↑↓) 3. <i>Consumption lost customers</i> (LRFM, ↓↓↑) 4. <i>Uncertain lost customers</i> (LRFM, ↓↓↓)
4	New Customers (NC)	Terdiri dari : 1. <i>High value new customers</i> (LRFM, ↓↑↑)
		2. <i>Frequency promotion customers</i> (LRFM, ↓↓↑) 3. <i>Spender promotion customers</i> (LRFM, ↓↓↓) 4. <i>Uncertain new customers</i> (LRFM, ↓↓↓)
5	Consuming Resource Customers (CRC)	Terdiri dari : 1. <i>Low consumption cost customers</i> (LRFM, ↑↓↓)
		2. <i>High consumption cost customers</i> (LRFM, ↑↑↓)

Hasil dari pemetaan pelanggan yang telah diperoleh tersebut dianalisa kembali untuk menentukan strategi *marketing* yang tepat untuk setiap segmen pelanggan.

Strategi *marketing* yang diusulkan kepada segmen pelanggan diperoleh berdasarkan *matrix* segmentasi strategi *marketing* dengan patokan nilai *frequency* dan *monetary* [21] yang dapat dilihat pada Tabel 3.

Tabel 3. *Matrix* Segmentasi Strategi *Marketing*

Nilai Pelanggan	<i>Monetary</i> Tinggi	<i>Monetary</i> Rendah
<i>Frequency</i> Tinggi	<i>Enforced</i> Strategi (Segmen I)	<i>Offensive</i> Strategi (Segmen II)
<i>Frequency</i> Rendah	<i>Defensive</i> Strategi (Segmen III)	<i>“Let-Go”</i> Strategi (Segmen IV)

3. Hasil dan Pembahasan

Data yang diperoleh merupakan rekapan data transaksi berupa excel sebanyak dua folder yang dimana masing-masing folder terdapat dua belas dokumen yang berisi rekapan transaksi harian dalam satu bulan. Data transaksi penyewaan bus tersebut memiliki sebanyak 14.292 data. Terdapat enam field pada mentah seperti tanggal yang menyatakan tanggal transaksi penyewaan, nomor polisi yang menyatakan nomor plat dari masing-masing bus, sopir yang menyatakan nama dari sopir yang mengendarai bus, agen/travel menyatakan nama pelanggan yang melakukan transaksi penyewaan, route menyatakan rute perjalanan yang ditempuh, dan harga menyatakan jumlah harga transaksi yang dilakukan dalam satuan rupiah. Dapat dilihat pada Tabel 4 yang merupakan beberapa contoh data mentah yang terdapat pada data transaksi, dimana digunakan sebagai tahapan dari persiapan data. Pembahasan adalah penjelasan dasar, hubungan dan generalisasi yang ditunjukkan oleh hasil. Uraianannya menjawab pertanyaan penelitian. Jika ada hasil yang meragukan maka tampilkan secara objektif.

Tabel 4. Data Mentah

Tgl	No. Polisi	Sopir	Agen/Travel	Route	Harga
1/1/17	DK 9141 CE	Arjiwa	Jan's Tour	Trf In to Pullman	380000
1/1/17	DK 9141 CF	Aditya	Maestro	Trf In to Grand Mirage	427500
1/1/17	DK 9141 DG	Defa	Matahari Korin	FD Golf	665000
1/1/17	DK 9141 DF	Widi	Ibu Surati Guide	FD	600000
...
30/12/18	57	Arjiwa	Bapak Yasa	FD	700000

Tahapan selanjutnya setelah mengumpulkan field data tersebut yaitu dilakukannya *preprocessing* data. Data yang telah diperoleh dari hasil saringan yaitu data yang berupa excel tersebut digabungkan ke dalam satu basis data SQL Server dengan cara mengimportkan data

transaksi yang dimana sebelumnya telah dikonversikan ke dalam file CSV kemudian dimasukkan ke dalam satu basis data yang telah dibuat. Perlunya menghilangkan atribut, catatan serta kolom yang tidak diperlukan pada proses reduksi data agar dapat menghindari volume data yang terlalu besar.

Dalam field harga terdapat *missing value* pada data transaksi yang bernilai nol. Pada penelitian ini field harga pada data transaksi yang terdapat *missing value* atau harga yang tercatat bernilai nol sebanyak 8 data dari total keseluruhan data yaitu 14.292 data transaksi. Total data yang terdapat *missing value* itu dapat dihilangkan karena jumlahnya hanya sedikit dan mempunyai atribut yang tidak bisa digunakan atau tidak bernilai untuk dilakukan pada proses selanjutnya. Atribut yang diperlukan dan yang tidak diperlukan dapat dilihat pada Tabel 5.

Tabel 5. Atribut yang dihilangkan dan yang digunakan

Atribut yang tidak diperlukan	Catatan yang dihilangkan	Atribut yang akan digunakan
No. Polisi	Harga yang tidak tertera/tercatat atau bernilai RP. 0	Tanggal
Sopir Route		Agen/travel Harga

Tahapan selanjutnya yaitu pembersihan data dimana data transaksi dengan atribut yang telah dipilih kemudian dilakukan pembersihan data khususnya pada nama pelanggan dengan mengambil beberapa karakter seperti tanda garis miring dan kata seperti driver, Guide, Camp, BT 676637, Drive, Travel Lombok, Tour Lombok, dan Beach Resort yang memiliki makna sama dan kata yang kurang lengkap. Pembersihan data dilakukan dengan bantuan bahasa pemrograman Python, dimana hasil dari pembersihan data tersebut kemudian dibuatkan tabel baru data transaksi dan telah terkoneksi langsung ke dalam basis data di SQL Server. Hasil pembersihan data dapat dilihat pada Tabel 6.

Tabel 6. Hasil Pembersihan Data

id_trans	tgl	nama_agen	Harga
1	1/1/2017	Ibu Bayu	2500000
2	1/1/2017	Ibu Komang	2600000
3	1/1/2017	Destination Asia	1300000
4	1/1/2017	Destination Asia	1300000
5	1/1/2017	Destination Asia	1300000
6	1/1/2017	Destination Asia	1300000
7	1/1/2017	Destination Asia	1300000
8	1/1/2017	Destination Asia	1300000
9	1/1/2017	Ibu Fatmawati	1500000
10	1/1/2017	Bapak Agung	1500000
...
14294	12/31/2018	PBI	475000

Dalam hal ini data transaksi yang diberikan tidaklah lengkap karena dalam proses selanjutnya diperlukannya atribut lain seperti ID *Customer*, maka dari itu perlu dibuatkan satu tabel lagi yaitu tabel pelanggan dan pengisian atribut ID *Customer* pada basis data di SQL Server. Hasil pembuatan ID *Customer* dengan bahasa pemrograman Python dan telah terkoneksi langsung pada basis data SQL Server dapat dilihat pada Tabel 7.

Tabel 7. Tabel *Customer*

id_cust	nama_agen
1	9 Travel
2	99 Travel
3	ABP Tour
4	ABW Tour
5	AJB Tour
6	AKBID Kartini
7	APH Tour
...	...
1821	the Champions Ubud

Data pelanggan dari perusahaan penyewaan bus ini kemudian dilakukan proses transformasi ke dalam masing – masing atribut LRFM (*Length, Recency, Frequency, Monetary*) untuk mengetahui jenis dan kelompok dari pelanggan tersebut. Proses transformasi ini dilakukan dengan *query* pada SQL Server. Nilai atribut *recency, length* dan *frequency* diperoleh dari tanggal transaksi yang dilakukan oleh pelanggan dan nilai *monetary* diperoleh dari variable harga pada data transaksi yang dimana merupakan total jumlah dalam satuan rupiah.

Perolehan nilai pada atribut *length* dan *recency* memiliki kesamaan karena batas untuk analisis periode transaksi yang ditetapkan adalah 1 Januari 2019 dan data transaksi yang digunakan pada penelitian yaitu selama dua tahun, dari Januari 2017 hingga Desember 2018. Dalam perhitungan *recency* dilakukan pencarian tanggal transaksi akhir dari pelanggan selama periode analisis dengan batas tanggal yang telah ditentukan, sedangkan untuk pencarian nilai *length* yaitu selisih periode waktu transaksi awal dan transaksi akhir oleh pelanggan selama batas periode analisis yang telah ditetapkan.

Hasil akhir yang dihasilkan pada nilai atribut *length* dan *recency* yaitu berupa satuan hari. Perolehan hasil transformasi data ke dalam masing – masing atribut LRFM dapat dilihat pada Tabel 8.

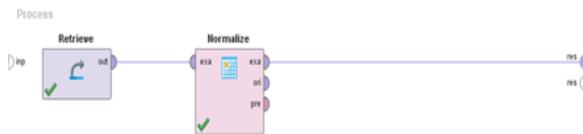
Tabel 8. Hasil Transformasi Data ke Nilai LRFM

id_cust	TGL	R	F	M	L
1	2017-07-17	-533	1	1300000	-533
2	2017-03-22	-650	1	1900000	-650
3	2017-08-31	-488	1	800000	-488
4	2017-09-01	-487	1	1200000	-487
5	2018-12-26	-6	1	550000	-6
6	2017-06-16	-564	3	3600000	-564
7	2017-04-28	-613	2	1350000	-613
7	2017-04-29	-612	1	1200000	-612
...
1821	2018-07-12	-25	1	2090000	-25

3.1. Tahapan *Data Mining* Metode *Clustering* K-Means

Sebelum masuk ke proses clustering, data yang sebelumnya telah melalui tahapan transformasi ke dalam masing – masing atribut LRFM (*Length, Recency, Frequency, Monetary*), selanjutnya yaitu menormalisasikan nilai LRFM. Normalisasi dilakukan karena melihat hasil nilai *monetary* yang memiliki rentang nilai cukup jauh dibandingkan dengan nilai *length, recency, dan frequency*.

Normalisasi menggunakan metode Z-score dengan bantuan *tools* RapidMiner. Proses normalisasi dapat dilihat pada Gambar 2.



Gambar 2. Proses Normalisasi dengan RapidMiner

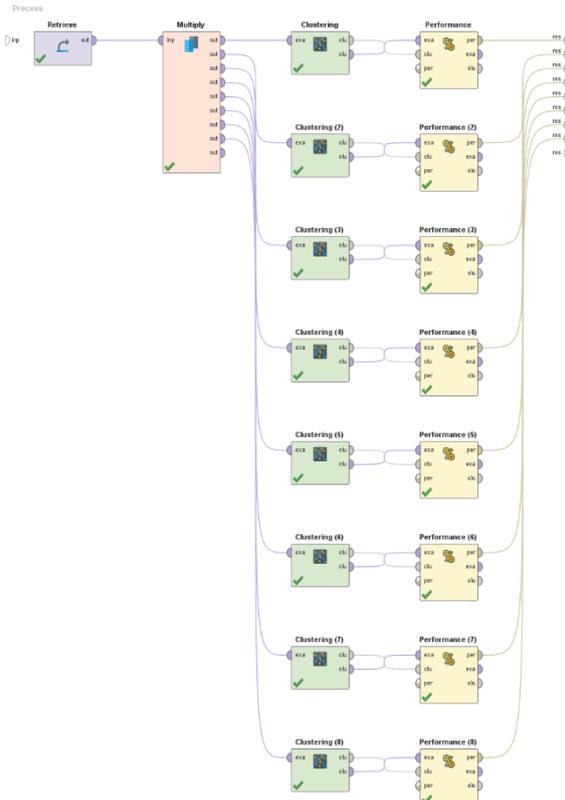
Proses normalisasi data dengan Z-score dari masing – masing atribut LRFM menghasilkan data yang siap untuk diolah pada proses metode *clustering*. Hasil normalisasi dapat dilihat pada Tabel 9.

Tabel 9. Hasil Normalisasi LRFM

Id_cst	Z_R	Z_F	Z_M	Z_L
1	-0.7611972	-0.41741835	-0.32657765	-0.7611972
2	-1.3126643	-0.41741835	-0.14362345	-1.3126643
3	-0.5490944	-0.41741835	-0.47903948	-0.5490944
4	-0.544381	-0.41741835	-0.35707002	-0.544381
5	1.72276173	-0.41741835	-0.5552704	1.7227617
6	-0.9073124	0.42859185	0.37474677	-0.9073124
7	-1.1382687	0.00558675	-0.31133147	-1.1382687
7	-1.1335553	-0.41741835	-0.35707002	-1.1335553
7	-1.1288419	-0.41741835	-0.35707002	-1.1288419
7	-1.1241285	-0.41741835	-0.35707002	-1.1241285
8	1.47295182	0.00558675	-0.02165399	1.4729518
9	0.52084614	0.00558675	0.03018304	0.5208461
9	0.18619514	0.42859185	0.37474677	0.1861951
9	0.92148467	-0.41741835	-0.1725912	0.9214847
...
1821	1.633207234	-0.41741835	-0.08568796	1.6332072

Data hasil normalisasi LRFM yang telah didapatkan sebelumnya dijadikan sebagai *dataset* untuk dilakukannya proses clustering. Sebelum masuk ke dalam pengolahan data dengan metode *clustering*, diperlukannya validasi jumlah *k-cluster* agar mendapatkan jumlah *cluster* yang optimal dan terbaik. Validasi jumlah *k-cluster* menggunakan *Davies Bouldin Index* (DBI) dengan bantuan *tools* RapidMiner. Proses validasi jumlah *cluster* dengan DBI dapat dilihat pada Gambar 3.

Pada proses ini, jumlah *k-cluster* diberikan nilai 3, 4, 5, 6, 7, 8, 9, dan 10 untuk diujicobakan agar mendapatkan jumlah *cluster* yang benar-benar optimal. Hasil yang diperoleh dari masing – masing perhitungan *performance* dengan *Davies Bouldin Index* (DBI) dapat dilihat pada Tabel 10.



Gambar 3. Proses Normalisasi dengan RapidMiner

Tabel 10. Hasil Davies Bouldin Index

Davies Bouldin Index	Jumlah Cluster
0.143	3
0.158	4
0.135	5
0.162	6
0.151	7
0.151	8
0.155	9
0.158	10

Perolehan jumlah *cluster* yang baik dan optimal untuk proses metode *clustering* selanjutnya adalah ditunjukkan pada penggunaan jumlah 5 *cluster* dengan nilai DBI yang minimal yaitu 0.135.

Tahapan selanjutnya adalah proses metode *clustering* dengan menggunakan K-Means. Proses pemodelan metode *clustering* dengan K-Means pada tool RapidMiner dapat dilihat pada Gambar 4.



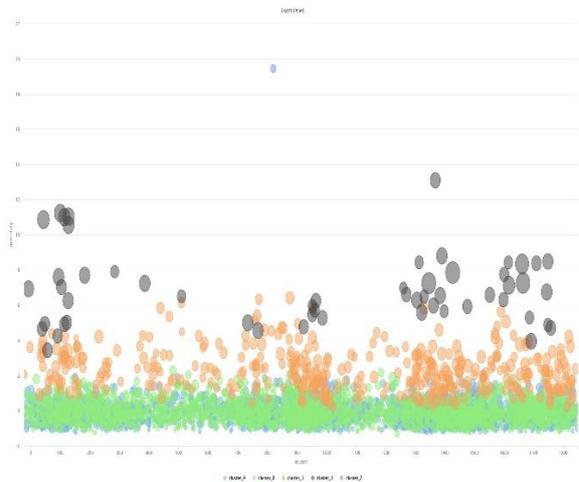
Gambar 4. Pemodelan Clustering K-Means dengan RapidMiner

Pada proses *clustering* dengan K-Means tersebut menghasilkan 5 *cluster* dari data transaksi pelanggan yang dapat dilihat pada Tabel 11.

Tabel 11. Hasil Clustering

id_cust	Cluster	R	F	M	L
5	0	1.723	-0.417	-0.56	1.723
8	0	1.473	0.0056	-0.02	1.473
...
1821	0	1.633	-0.417	-0.09	1.633
47	1	-0.7	7.6197	6.87	-0.7
76	1	1.002	10.158	11.21	1.002
...
1801	1	-1.053	5.9277	4.613	-1.053
875	2	-0.045	0.4286	19.7	-0.045
9	3	0.318	0.8516	1.479	0.318
...
1818	3	1.157	3.3896	1.716	1.157
1	4	-0.761	-0.417	-0.33	-0.761
2	4	-1.313	-0.417	-0.14	-1.313
...
1818	4	-0.219	0.4286	0.375	-0.219

Hubungan antara nilai atribut *frequency*, *monetary*, ID *Customer*, dan setiap *cluster* dapat dilihat pada Gambar 5.



Gambar 5. Visualisasi Cluster dengan RapidMiner

Frequency dan *monetary* memiliki hubungan yang kuat dalam pembagain data ke dalam masing – masing *cluster*. Data yang berada pada *cluster* yang sama mempunyai sifat kemiripannya antara satu dengan yang lain pada nilai *frequency* dan *monetary*.

3.2. Analisis Hasil Clustering

Hasil dari proses *clustering* yang telah dilakukan kemudian dianalisa sebagai tahapan untuk dapat mencari jenis kelompok pelanggan berdasarkan model LRFM. Perhitungan standar deviasi dilakukan dalam pencarian LRFM index untuk mengetahui karakteristik pelanggan dengan mengacu kepada *customer value matrix* dan *customer loyalty matrix*. Perhitungan standar deviasi dilakukan dengan bahasa pemrograman Python. Hasil perhitungan standar deviasi dapat dilihat pada Tabel 12.

Tabel 12. Hasil Perhitungan Standar Deviasi LRFM

Cluster	L	R	F	M	Simbol LRFM
0	0.464	0.464	0.389	0.426	L↓R↓F↓M↓
1	0.884	0.884	3.09	2.1	L↑R↑F↑M↑
2	0	0	0	0	L↓R↓F↓M↓
3	0.853	0.853	1.068	1.121	L↑R↑F↑M↑
4	0.478	0.478	0.387	0.378	L↓R↓F↓M↓
Rata rata	0.536	0.536	0.987	0.805	

Simbol LRFM didapat dari hasil perhitungan standar deviasi atribut LRFM dengan rata – rata dari setiap *cluster*. Simbol tanda panah ke atas (↑) dihasilkan jika hasil standar deviasi dari atribut LRFM lebih tinggi dari rata – ratanya, sedangkan untuk simbol tanda panah ke bawah (↓) dihasilkan jika hasil standar deviasi dari atribut LRFM lebih rendah dari rata – ratanya. Hasil pengelompokan pelanggan dapat dilihat pada Tabel 13.

Tabel 13. Penggolongan Kelompok Pelanggan

Cluster	Kelompok Pelanggan	Jenis Pelanggan	LRFM Index
0	New Customers	Uncertain new customers	L↓R↓F↓M↓
1	Potential Customers	Potential loyal customers	L↑R↑F↑M↑
2	New Customers	Uncertain new customers	L↓R↓F↓M↓
3	Potential Customers	Potential loyal customers	L↑R↑F↑M↑
4	New Customers	Uncertain new customers	L↓R↓F↓M↓

Pada tabel di atas *cluster 0*, *cluster 2*, dan *cluster 4* merupakan kelompok pelanggan baru yang keberadaannya tidak pasti dalam meningkatkan keuntungan bagi perusahaan.

Cluster 1 dan *cluster 3* merupakan kelompok pelanggan yang berpotensi setia dan loyal terhadap perusahaan dalam melakukan transaksi.

Berdasarkan hasil analisis *cluster* dan *index* LRFM yang telah dilakukan terdapat perbedaan sikap pembelian atau transaksi pada pelanggan yang dimana hal tersebut memberikan keuntungan yang beragam bagi perusahaan, maka dari itu perlunya strategi *marketing* dalam mengoptimalkan keuntungan bagi perusahaan. Strategi *marketing* yang diusulkan bagi setiap *cluster* yang telah terbentuk yaitu menggunakan *matrix* segmentasi strategi *marketing* dengan patokan nilai *frequency* dan *monetary*.

Usulan strategi *marketing* bagi pelanggan yang berada pada *cluster 0*, *cluster 2*, dan *cluster 4* dengan jenis pelanggan adalah *uncertain new customers* yaitu *let-go* strategi. *Let go* strategi ini dipilih karena melihat dari nilai *index* LRFM yang dihasilkan. (L↓ R↓ F↓ M↓),

dimana jarang nya melakukan transaksi dan jumlah transaksi yang dikeluarkan relatif sedikit yang dapat dilihat dari nilai *frequency* dan *monetary*. Segmen pelanggan ini merupakan termasuk pelanggan baru dan tidak pasti. Cara yang dapat dilakukan bagi perusahaan untuk bisa membuat pelanggan pada segmen ini menjadi lebih loyal yaitu seperti menjalin komunikasi yang kontinu dengan para pelanggan untuk bisa mengetahui kebutuhan yang dibutuhkan, adapun juga seperti memberikan produk yang berfokus pada keinginan para pelanggan yang dimana kedepannya pelanggan tersebut dapat dikelola sampai menjadi pelanggan yang loyal dan berpotensi menguntungkan bagi perusahaan.

Usulan strategi *marketing* bagi pelanggan yang berada pada *cluster 1* dan *cluster 3* dengan jenis pelanggan adalah *potential loyal customers* yaitu *enforced* strategi. Strategi ini dipilih karena ditimbang dari LRFM *index* (L↑ R↑ F↑ M↑) yang dihasilkan. *Enforced* strategi ini digambarkan dari ciri khas para pelanggan yang mempunyai nilai transaksi cukup sering dengan perusahaan, jumlah transaksi yang dilakukan cukup tinggi dan juga sangat berkontribusi besar bagi perusahaan karena dapat dilihat dari *frequency* dan *monetary* yang tinggi. Salah satu cara yang bisa diimplementasikan oleh perusahaan seperti memberikan informasi mengenai promosi pada hari – hari tertentu atau promosi yang sesuai pada kebutuhan setiap pelanggan. Perusahaan perlu memahami dan mengerti kebutuhan maupun perilaku dari setiap pelanggan yang ada pada kategori ini untuk bisa memberikan pelayanan yang baik dan optimal. Perusahaan juga dapat meningkatkan pelayanan kepada pelanggan agar pelanggan merasa nyaman dan transaksi yang dilakukan dapat bertahan lama.

4. Kesimpulan

Proses implementasi *clustering* dilakukan dengan beberapa tahapan yaitu *preprocessing* data, transformasi data ke dalam atribut LRFM, penentuan *cluster* dengan Davies Bouldin *Index*, dan implementasi Algoritma K-Means. Pemetaan data ke dalam model LRFM menyesuaikan dengan atribut dari LRFM (*Length*, *Recency*, *Frequency*, dan *Monetary*). Proses validasi dengan menggunakan *index* validitas Davies Bouldin *Index* dengan parameter jumlah *cluster* dari tiga hingga sepuluh tersebut menunjukkan bahwa jumlah *cluster* terbaik untuk mendapatkan hasil yang optimal yaitu dengan perolehan nilai Davies Bouldin *Index* 0.135 pada jumlah *cluster* yaitu 5. Berdasarkan hasil *clustering* tersebut dilakukanlah analisis untuk mengetahui jenis kelompok setiap pelanggan dengan menggunakan perhitungan standar deviasi. Hasil perhitungan standar deviasi dari 5 *cluster* dengan masing – masing atribut LRFM didapatkan bahwa *cluster 0*, *cluster 2*, dan *cluster 4* termasuk kelompok pelanggan *new customers* dengan jenis *uncertain new*

customers (L↓R↓F↓M↓) dan strategi *marketing* yang diusulkan yaitu *let-go* strategi. Cluster 1 dan cluster 3 termasuk kelompok pelanggan *potential customers* dengan jenis *potential loyal customers* (L↑R↑F↑M↑) dan strategi *marketing* yang diusulkan yaitu *enforced* strategi. Hasil yang telah diperoleh ini dapat membantu perusahaan dalam mempertahankan pelanggan yang berpotensi menguntungkan bagi perusahaan kedepannya.

Saran yang ditujukan untuk penelitian selanjutnya yaitu dapat dilakukan perbandingan dalam metode *clustering* dengan jenis metode *clustering* lainnya untuk mengetahui tingkat keakuratan yang lebih tinggi dalam mendapatkan hasil *clustering* yang lebih baik.

Daftar Rujukan

- [1] P. Wahyu Adityo, "Sepanjang 2019, Kunjungan Wisman ke Bali dari Sejumlah Negara Meningkat," *10/10/2019*, 2019. <https://travel.kompas.com/read/2019/10/10/211500427/sepanjang-2019-kunjungan-wisman-ke-bali-dari-sejumlah-negara-meningkat>. (diakses Jul 03, 2021).
- [2] "Maibus." <https://www.maibustransport.com/> (diakses Jun 30, 2021).
- [3] Y. O. Rahmawati, A. Kusniawati, dan I. Setiawan, "Pengaruh Customer Relationship Management dan Kualitas Pelayanan Terhadap Loyalitas Konsumen Sepeda Motor Yamaha (Studi pada Konsumen Bahana Ciamis)," *Bus. Manag. Entrep. J.*, vol. 1, no. 4, hal. 102–115, 2019.
- [4] S. Eka, "Segmentasi Pelanggan Menggunakan Two Stage Clustering dan LRFM Model pada Divisi Marketing PT . XYZ untuk Mendukung Strategi Pengelolaan Pelanggan," hal. 1–7, 2011.
- [5] T. Hardiani, S. Sulistyono, dan R. Hartanto, "Kajian Data Mining Customer Relationship Management pada Lembaga Keuangan Mikro," *J. Sist. Inf. Bisnis*, vol. 4, no. 3, hal. 181–187, 2014, doi: 10.21456/vol4iss3pp181-186.
- [6] M. F. Mulya, N. Rismawati, dan A. R. Rizky, "Analisis Dan Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Penjualan Pada Kantin Universitas Tanri Abeng," *Fakt. Exacta*, vol. 12, no. 3, hal. 210–218, 2019, doi: 10.30998/faktorexacta.v12i3.4541.
- [7] Y. H. Chrisnanto dan A. Kaniangsih, "Pengelompokan Ekuitas Pelanggan Berbasis Recency Frequency Monetary (Rfm) Menggunakan K-Means Clustering," *Semin. Nas. Teknol. Inf. dan Komun. 2019 (SENTIKA 2019)*, vol. 2019, no. Sentika, hal. 13–14, 2019, [Daring]. Tersedia pada: <https://fti.uajy.ac.id/sentika/publikasi/makalah/2019/1.pdf>
- [8] D. P. Hidayatullah, R. I. Rokhmawati, dan A. R. Perdanakusuma, "Analisis Pemetaan Pelanggan Potensial Menggunakan Algoritma K-Means dan LRFM Model Untuk Mendukung Strategi Pengelolaan Pelanggan (Studi Pada Maninjau Center Kota Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 8, hal. 2406–2415, 2018.
- [9] S. Monalisa, "Klusterisasi Customer Lifetime Value dengan Model LRFM menggunakan Algoritma K-Means," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 2, hal. 247, 2018, doi: 10.25126/jtiik.201852690.
- [10] F. Marisa, S. S. S. Ahmad, Z. I. M. Yusof, Fachrudin, dan T. M. A. Aziz, "Segmentation model of customer lifetime value in Small and Medium Enterprise (SMEs) using K-Means Clustering and LRFM model," *Int. J. Integr. Eng.*, vol. 11, no. 3, hal. 169–180, 2019, doi: 10.30880/ijie.2019.11.03.018.
- [11] F. Hadi, M. Mustakim, D. O. Rahmadia, F. H. Nugraha, N. P. Bulan, dan S. Monalisa, "Penerapan K-Means Clustering Berdasarkan RFM Mofek Sebagai Pemetaan dan Pendukung Strategi Pengelolaan Pelanggan (Studi Kasus: PT. Herbal Penawar Alwahidah Indonesia Pekanbaru)," *J. Sains dan Teknol. Ind.*, vol. 15, no. 1, hal. 69–76, 2017.
- [12] B. Rahmat C.T.I. et al., "Implementasi k-means clustering pada rapidminer untuk analisis daerah rawan kecelakaan," *Semin. Nas. Ris. Kuantitatif Terap. 2017*, no. April, hal. 58–60, 2017.
- [13] M. T. Dharmawan, N. Y. Setiawan, dan F. A. Bachtiar, "Segmentasi Pelanggan Menggunakan Metode Fuzzy C-Means Clustering Berdasarkan LRFM Model Pada Toko Sepatu (Studi Kasus : Ride Inc Kota Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 2, hal. 1978–1985, 2019.
- [14] I. D. A. A. Y. Primandari, I. K. G. D. Putra, dan I. M. Sukarsa, "Customer Segmentation Using Particle Swarm Optimization and K-Means Algorithm," *Int. J. Digit. Content Technol. its Appl.*, vol. 10, no. 4, hal. 22–28, 2016.
- [15] N. P. P. Yuliani, I. K. G. D. Putra, dan N. K. D. Rusjayanti, "Customer segmentation through fuzzy C-means and fuzzy RFM method," *J. Theor. Appl. Inf. Technol.*, vol. 78, no. 3, hal. 380–385, 2015.
- [16] D. A. Nasution, H. H. Khotimah, dan N. Chamidah, "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, hal. 78, 2019, doi: 10.24114/cess.v4i1.11458.
- [17] H. A. Prihanditya, "The Implementation of Z-Score Normalization and Boosting Techniques to Increase Accuracy of C4.5 Algorithm in Diagnosing Chronic K
- [18] M. Using dan D. Bouldin, "Optimalisasi Jumlah Kelompok Pada Metode Single Linkage," 2016.
- [19] G. Haris, P. Wibawa, I. G. Made, A. Sasmita, dan I. M. S. Raharja, "Analisis Data Log HoneyPot Menggunakan Metode K-Means Clustering," *J. Ilm. Merpati Univ. Udayana*, vol. 8, no. 1, hal. 13–21, 2020.
- [20] A. Rohmah, F. Sembiring, dan ..., "Implementasi Algoritma K-Means Clustering Analisis Untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: Smk Yaspim ...)," ... *Sist. Inf. dan ...*, hal. 290–298, 2021, [Daring]. Tersedia pada: <https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/32>.
- [21] N. Lubis, "Penerapan Customer Relationship Management (Crm) Dengan Menggunakan Metode Lrfm Analysis," *J. Din. Manaj. Dan Bisnis*, vol. 1, no. 2, hal. 283, 2018, doi: 10.21009/jdmb.01.2.06.