

IEEE Journal

by FATHORAZI NUR FAJRI

Submission date: 25-Sep-2019 10:50AM (UTC+0700)

Submission ID: 768516899

File name: Sumber.docx (136.33K)

Word count: 3202

Character count: 19211

Analisis RFM dan Algoritma K-Means untuk Clustering Loyalitas Customer

Jamal, S.E., M.Kom.¹⁾, Dwi Yanto, S.E., M.Kom.²⁾
¹⁾Dosen Komputerisasi Akuntansi - AMIK Taruna Probolinggo
Jl. Raya Leles No. A3, Leles Probolinggo
Telp : (0335) 681497
E-mail : tanpade@gmail.com¹⁾

ABSTRAK

Bagi sebuah perusahaan loyalitas customer menjadi hal yang sangat penting untuk diketahui. Pemahaman yang baik terhadap customer dapat digunakan untuk menentukan customer yang potensial atau pun yang tidak potensial. Customer potensial adalah customer yang memiliki tingkat kepuasan yang tinggi dengan frekwensi dan nilai permintaan yang tinggi pula. Kepuasan customer dapat terjadi apabila harapan, kebutuhan dan keinginan customer dapat dipenuhi sehingga akan selalu melakukan pembelian secara terus menerus yang pada akhirnya meningkatkan omzet perusahaan. Pada umumnya untuk menghasilkan clustering yang baik metode yang sering digunakan adalah metode analisis RFM (*Recency, Frequency, Monetary*). RFM sendiri adalah sebuah metode analisis yang cukup terkenal dan sering digunakan dalam strategi pemasaran melalui pendekatan CRM (*Customer Relationship Management*). RFM menghendaki penilaian loyalitas tidak hanya dipandang dari besarnya nilai pembelian (*monetary*) yang dilakukan oleh customer tetapi juga melibatkan tingkat keseringan (*frequency*) dan waktu terakhir (*recency*) seorang customer melakukan transaksi. Dataset yang dihasilkan dikelompokkan (*clustering*) menggunakan algoritma *K-Means* dengan 5 tingkat loyalitas yakni Sangat Loyal, Cukup Loyal, Biasa, Kurang Loyal, Tidak Loyal. Untuk uji validitas digunakan *Davies-Bouldin Index*. Nilai DBI yang dihasilkan dari clustering customer sebesar 0.79074. Dari nilai DBI tersebut dapat disimpulkan bahwa kualitas cluster yang dihasilkan memiliki kualitas yang cukup baik.

Kata kunci: *Clustering, Loyalitas, RFM, K-Means, DBI*

ABSTRACT

For a company customer loyalty is very important to know. A good understanding of customers can be used to determine potential or non-potential customers. Potential customers are customers who have a high level of satisfaction with a high frequency and value of requests as well. Customer satisfaction can occur if the expectations, needs and desires of customers can be met so that they will always make purchases continuously which ultimately increases the company's turnover. In general, to produce a good clustering method that is often used is the RFM (*Recency, Frequency, Monetary*) analysis method. RFM itself is a method of analysis that is quite well known and is often used in marketing strategies through the CRM (*Customer Relationship Management approach*). RFM requires that the assessment of loyalty is not only seen from the amount of purchase value (*monetary*) made by the customer but also involves the level of frequency (*frequency*) and the last time (*recency*) a customer makes a transaction. The resulting dataset is grouped (*clustering*) using the *K-Means* algorithm with 5 levels of loyalty namely Very Loyal, Fairly Loyal, Ordinary, Less Loyal, Not Loyal. For the validity test the *Davies-Bouldin Index* is used. The DBI value generated from customer clustering is 0.79074. From the DBI value it can be concluded that the quality of the cluster produced has quite good quality.

Keywords: *Clustering, Loyalty, RFM, K-Means, DBI*

I. PENDAHULUAN

I.1. Latar Belakang Masalah

Setiap perusahaan berkepentingan untuk mengetahui tingkat loyalitas customernya. Dengan mengetahui tingkat loyalitas customer perusahaan dapat menentukan langkah terbaik untuk meningkatkan omzet dengan memelihara hubungan baik terhadap customer misalnya memberikan perlakuan khusus kepada customer seperti memberikan souvenir atau hadiah setiap tahun yang besarnya ditentukan berdasarkan tingkat loyalitasnya. Customer menduduki posisi sangat penting dalam pengembangan strategi perusahaan, customer juga merupakan bagian dari sumber keuntungan perusahaan. Untuk itu diperlukan pemahaman yang baik mengenai customer [1]. Pemahaman yang baik tentang customer dapat digunakan untuk menentukan customer yang potensial ataupun yang tidak potensial. Customer yang potensial adalah pelanggan yang memiliki tingkat kepuasan yang tinggi sehingga memiliki loyalitas yang tinggi dengan frekwensi dan nilai permintaan yang tinggi pula [2]. Kepuasan customer dapat terjadi apabila harapan, kebutuhan dan keinginan customer dapat dipenuhi sehingga akan selalu melakukan pembelian secara terus menerus yang pada akhirnya meningkatkan omzet perusahaan [3]. Demikian pula sebaliknya, customer yang tidak puas akan mengembalikan barang yang dibeli kepada perusahaan atau dalam istilah bisnis disebut dengan retur penjualan. Retur penjualan adalah aktifitas transaksi yang mengurangi nilai penjualan [4]. Transaksi retur penjualan terjadi jika perusahaan menerima pengembalian barang dari customer [5]. Semakin tinggi nilai retur terhadap penjualan maka kepuasan customer semakin menurun dan hal tersebut akan menurunkan loyalitas customer [6].

Untuk meningkatkan kualitas hubungan dengan customer umumnya perusahaan menerapkan suatu pendekatan yang disebut dengan CRM (*Customer Relationship Management*). CRM adalah suatu pendekatan untuk memahami customer, mempengaruhi perilaku dengan komunikasi yang efektif untuk menyaring customer baru, menjaga masa customer, loyalitas dan profitabilitas customer [7].

Pada data transaksi customer umumnya memiliki banyak atribut dan ini bisa menyebabkan pengolahan data nantinya kurang baik. Untuk itu perlu adanya pemilihan atribut-atribut yang memiliki pengaruh kuat terhadap tujuan pengolahan data sehingga hasilnya lebih optimal. Untuk mengatasi masalah pemilihan atribut terbaik diusulkan menggunakan metode RFM. RFM (*Recency Frequency Monetary*) adalah metode yang umum digunakan untuk mengoptimalkan data loyalitas customer yaitu dengan membagi aktifitas customer ke dalam 3 atribut yakni *recency*, *frequency* dan *monetary*. Dengan metode RFM data customer dititikberatkan pada 3 atribut yakni *recency*, *frequency* dan *monetary* [8]. Masing-masing atribut mewakili konversi data ke dalam 5 group kuartil yang sama. Atribut *Recency* adalah jarak terdekat terhadap waktu diambilnya data. Waktu di sini bisa dalam satuan hari, bulan maupun tahun. Dalam penelitian ini digunakan satuan hari. Semakin dekat dengan waktu pengambilan data, maka semakin tinggi poinnya. Atribut *Frequency* mewakili jumlah transaksi yang terjadi selama periode waktu tertentu. Semakin tinggi nilainya menunjukkan semakin sering customer melakukan transaksi, semakin potensial customer tersebut dan semakin tinggi poinnya. Sedangkan atribut *Monetary* untuk mewakili total nilai transaksi selama periode tertentu. Semakin tinggi nilai transaksi semakin tinggi poin *monetary*-nya.

Di samping memiliki kelebihan RFM juga memiliki beberapa kekurangan antara lain RFM berfokus hanya pada customer terbaik, memiliki variabel yang terbatas, berfokus hanya pada customer yang ada dan tidak dapat diterapkan pada customer baru atau calon customer, RFM juga melakukan respon yang sama terhadap semua data customer padahal pada realitanya setiap customer memiliki karakteristik yang berbeda [9].

Untuk mengelompokkan customer diperlukan suatu metode pengelompokan atau *clustering*. Saat ini telah banyak algoritma yang digunakan untuk segmentasi customer seperti metode SOM (*Self Organizing Map*), *Fuzzy C-Means* dan *K-Means*. Metode Algoritma *K-Means* adalah metode *Clustering* yang paling populer digunakan karena beberapa kelebihan yang

dimiliki antara lain algoritma ini sederhana dan mudah diterapkan [1]. Algoritma *Fuzzy C-Means* lebih digunakan untuk dataset dengan atribut yang banyak, sedangkan *K-Means* lebih digunakan untuk dataset dengan atribut yang sedikit [10]. Salah satu kekurangan dari *K-Means* adalah tidak adanya ketentuan yang pasti berapa jumlah partisi *cluster* (*K*) yang terbaik sedangkan penentuan jumlah partisi yang berbeda dapat menghasilkan keanggotaan yang berbeda pula [1].

Demikian pula lah yang dialami UD Sumber yang berlokasi di Grati Pasuruan. Setiap tahun perusahaan distributor sembako dan kebutuhan rumah tangga ini membagikan hadiah kepada customer dengan tujuan mengikat loyalitas customernya. Selama ini pembagian hadiah hanya didasarkan kepada ranking besarnya omzet yang dihasilkan dari setiap customer. Dengan jumlah customer yang cukup besar mencapai 503 customer potensial berupa grosir dan retail yang tersebar di berbagai wilayah di 4 kabupaten yakni Kabupaten dan Kota Pasuruan, Kabupaten dan Kota Probolinggo maka pengolahan customer menjadi masalah tersendiri. Dengan memanfaatkan teknik analisis RFM dan algoritma *clustering K-Means* serta uji validasi cluster diharapkan perusahaan dapat dengan mudah mengelompokkan customer. Dengan demikian perusahaan akan lebih mudah mengambil langkah strategi manajemen customer yang pada akhirnya meningkatkan omzet perusahaan.

16 1.2. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengelompokkan customer ke dalam 5 (lima) tingkat loyalitas menggunakan suatu metode yang umum digunakan yakni metode analisis RFM dan algoritma *clustering K-Means*. Pembagian loyalitas customer dikelompokkan menjadi: Sangat Loyal, Cukup Loyal, Biasa, Kurang Loyal, dan Tidak Loyal. Kelompok customer yang dihasilkan harus mampu diuji tingkat kualitasnya. Untuk itu diperlukan suatu alat validasi yang mampu mengukur kualitas dari setiap kelompok customer tersebut. Analisa kualitas cluster yang digunakan adalah DBI (*Davies Bouldin Index*). Dengan diperolehnya pengelompokan ini maka perusahaan dapat menentukan strategi

manajemen yang tepat dalam memelihara loyalitas customer kepada perusahaan. Dengan demikian omzet perusahaan pada akhirnya lebih meningkat.

1.3. Penelitian Terkait

Banyak sekali penelitian tentang kuatnya pengaruh kepuasan terhadap loyalitas customer. Salah satu penelitian dilakukan Prof. Chandrashekar et. al. [6] terhadap perusahaan ABC di Amerika Serikat yang melayani ratusan perusahaan besar yang beroperasi di bisnis-bisnis layanan di Amerika Serikat dan Kanada. Pada penelitian tersebut, penulis fokus pada identifikasi penyebab rentannya customer berpindah karena masalah kepuasan mereka. Model yang digunakan adalah Judgment Uncertainty and Magnitude Parameters (*JUMP*) model, penulis mengurai kepuasan customer menjadi dua sisi yakni aspek kepuasan dan tingkat kepuasan.

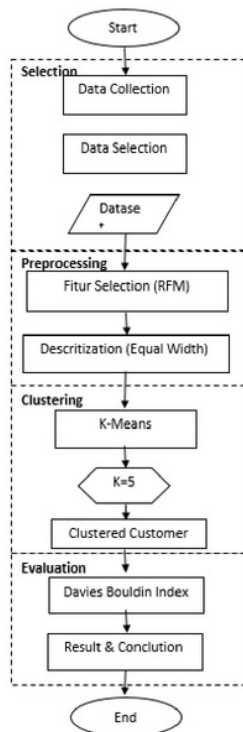
Pengukuran akhirnya adalah peran kepuasan terhadap loyalitas customer. Hasil studi menunjukkan bahwa kekuatan kepuasan memainkan peran sentral dalam tingkat loyalitas customer. Di mana hasilnya adalah bahwa ketidakpuasan sangat mempengaruhi loyalitas secara signifikan yakni menurunkan rata-rata sekitar 60%. Penelitian juga menunjukkan bahwa aspek hubungan sebelumnya (panjang hubungan, volume bisnis, dan kesukaan dari pengalaman sebelumnya) mengakibatkan kerentanan lebih besar.

Penelitian lainnya dilakukan oleh Hosseini & Maleki, 2010 [11]. Penelitian ini dilakukan pada perusahaan SAPCO Corp. Iran merupakan salah satu perusahaan pemasok paling terkemuka di Iran yang menyediakan barang-barang, peralatan dan aksesoris mobil untuk beberapa pabrik mobil di Iran melalui beberapa perusahaan agen distributor besar seperti ISACO Corp. dengan pelayanan prima jual yang maksimal bagi para customer.

Pada penelitian ini peneliti juga menggunakan analisis RFM dan Algoritma *K-Means* namun juga memasukkan satu parameter baru pada RFM yakni Length of relationship (disimbolkan dengan *L*) yang dihitung dari pertama kali customer bergabung pada perusahaan. Untuk optimalisasi digunakan *Davies-Boildin* Indeks, mengelompokkan loyalitas customer

produk di bawah konsep B2B. Hasil penelitian adalah, dengan determinasi optimum (DM) tanpa L diperoleh nilai DBI 0.186094, dengan melibatkan L diperoleh nilai DBI 0.181831. Metodologi yang dikembangkan telah diimplementasikan pada perusahaan tersebut. Hasilnya menunjukkan kemampuan luar biasa di mana perusahaan memiliki kemampuan menilai loyalitas customer lebih baik. Perusahaan juga mampu mengatur strategi pemasaran yang lebih dibandingkan dengan perusahaan besar lainnya di Iran.

2. METODE



Gambar 1. Metode yang Dilakukan

Sebagaimana ditunjukkan pada gambar 1 metode yang akan diterapkan dalam penelitian ini meliputi:

1. Selection

Dataset awal diambil dari data penjualan yang terjadi selama tahun 2018. Dari beberapa tabel penjualan yang ada dikumpulkan (*collection*) untuk selanjutnya dipilih (*selection*) dan digabungkan data-data yang diperlukan untuk penelitian ini (*selected data*).

2. Preprocessing

Tahap ini digunakan untuk mempersiapkan dataset agar memiliki kualitas yang lebih baik dan efektif sebelum dimodelkan. Pada tahap ini dilakukan seleksi fitur menggunakan metode RFM (*fitur selection by RFM*), dihasilkan 3 atribut yakni *recency*, *frequency*, dan *monetary*. Dilanjutkan tahap *descritization* data dengan teknik *equal width* yakni memberikan bobot pada setiap fitur data kontinyu sehingga menjadi data diskrit. Dengan demikian diperoleh dataset customer dengan fitur yang sederhana dengan atribut: *kdcustomer*, *recency*, *frequency*, *monetary*.

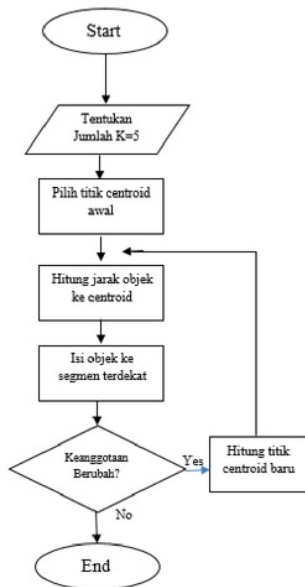
3. Clustering

Pengklasteran dataset customer digunakan *K-Means* dengan nilai $K=5$. Penentuan nilai $K=5$ adalah untuk mengelompokkan customer yang sesuai dengan pembobotan yang dilakukan oleh teknik RFM yakni pembobotan customer dikategorikan dengan nilai 5, 4, 3, 2, 1 dengan konversi 5 = "Sangat Loyal", 4 = "Loyal", 3 = "Biasa", 2 = "Kurang Loyal", 1 = "Tidak Loyal". Sebagai konsekwensinya maka nilai koordinat centroid awal ditentukan seperti tabel 1.

Tabel 1. Penentuan Centroid Awal

| Centroid | Nilai awal | Keterangan |
|----------|------------|--------------|
| C1 | 1,1,1 | Tidak Loyal |
| C2 | 2,2,2 | Kurang Loyal |
| C3 | 3,3,3 | Biasa |
| C4 | 4,4,4 | Loyal |
| C5 | 5,5,5 | Sangat Loyal |

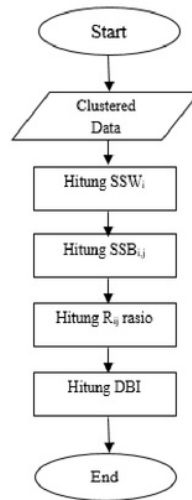
Langka-langkah clustering dengan K-Means adalah sebagai berikut:



Gambar 2. Langkah K-Means

4. Evaluation

Tahap selanjutnya adalah pengukuran atau evaluasi. Pengukuran *Index Davies Bouldin (DBI)* adalah memaksimalkan jarak antar *cluster* di antara C_i dan C_j dan pada waktu yang sama mencoba untuk meminimalkan jarak antar titik dalam sebuah *cluster*. Langkah-langkah evaluasi *cluster* yang akan dilakukan terhadap data RFM menggunakan *Davies Bouldin Index* adalah sebagai berikut:



Gambar 3. Langkah Evaluasi Cluster dengan DBI

Berdasarkan *Indeks Davies-Bouldin, clustering* yang optimal adalah yang memiliki *Indeks Davies-Bouldin* minimal. Semakin kecil nilai *SSW* maka hasil *clustering* yang didapat adalah semakin baik. Sedangkan pada *SSB* semakin besar nilainya maka semakin baik. Karena hal tersebut menunjukkan semakin kuatnya perbedaan atau jarak antar *cluster*. Demikian halnya dengan nilai *DBI*, secara esensial *DBI* menginginkan nilai sekecil mungkin (non-negatif ≥ 0) untuk menilai baiknya *cluster* yang didapat. *DBI* seringkali juga digunakan untuk menentukan jumlah *K* yang paling ideal yang digunakan pada algoritma *K-Means* [12].

3. HASIL DAN PEMBAHASAN

pembobotan ditunjukkan sebagaimana pada tabel 3.

Field Code Changed

3.1. Tahap Selection dan Preprocessing

Data transaksi penjualan diproses melalui tahap *selection (collection, selection)* dengan melakukan rekap penjualan setiap customer. Selanjutnya dengan proses *preprocessing* dengan metode RFM dimana pada tahap *Fitur Selection* didapat variable *Recency, Frequency* dan *Monetary (RFM)* sebagai berikut:

Tabel 2. Dataset RFM

| KdCus | R | F | M |
|----------|-----|-----|------------|
| BGL | 17 | 6 | 12.304,369 |
| BGL001 | 11 | 17 | 4.460,224 |
| BGL10000 | 255 | 1 | 119,790 |
| BGL1310 | 10 | 15 | 17.707,137 |
| BGL965 | 3 | 88 | 84.294,878 |
| BGL001 | 24 | 18 | 90.426,116 |
| BGL0011 | 311 | 2 | 16.242,360 |
| BGL0012 | 59 | 12 | 47.325,193 |
| BGL002 | 3 | 24 | 18.677,775 |
| BGL003 | 3 | 15 | 7.145,190 |
| BGL004 | 31 | 17 | 4.687,467 |
| BGL005 | 11 | 31 | 10.915,728 |
| BGL01 | 23 | 17 | 56.688,000 |
| BGL011 | 3 | 25 | 11.678,646 |
| BGL02 | 11 | 44 | 17.636,057 |
| BGL04 | 3 | 26 | 12.558,696 |
| BGL05 | 3 | 68 | 68.258,397 |
| BGL06 | 3 | 31 | 10.564,698 |
| BGL09 | 101 | 18 | 15.901,619 |
| BGL1 | 31 | 20 | 6.738,040 |
| ... | ... | ... | ... |

Selanjutnya pada proses *preprocessing* ini dilakukan tahap *discretization* yakni proses memberikan bobot pada setiap fitur data kontinyu sehingga menjadi data diskrit kategorikal. Dengan demikian diperoleh dataset customer dengan fitur yang sederhana. Dari indek rasio data dibagi ke dalam 5 bagian dalam pembobotan 5,4,3,2,1 dengan konversi 5="Sangat Puas", 4="Puas", 3="Biasa", 2="Kurang Puas", 1="Tidak Puas". Bagian dengan nilai rasio terbesar diberi bobot 5 karena menandakan customer semakin puas. Demikian seterusnya sampai customer dengan nilai rasio terkecil diberi bobot 1. Hasil pembobotan ini menjadi dataset dasar untuk diolah pada pengolahan selanjutnya yakni pengelompokan customer dengan menggunakan K-Means. Interval data

Tabel 3. Interval Pembobotan

| Recency | Dari | Sampai |
|-----------|-------------|-------------|
| 5 | 0 | 18 |
| 4 | 19 | 36 |
| 3 | 37 | 54 |
| 2 | 55 | 72 |
| 1 | 73 | 90 |
| Frequency | Dari | Sampai |
| 1 | 1 | 19 |
| 2 | 20 | 37 |
| 3 | 38 | 56 |
| 4 | 57 | 74 |
| 5 | 75 | 93 |
| Monetary | Dari | Sampai |
| 1 | 5.003,055 | 62.558,264 |
| 2 | 62.558,265 | 120.113,473 |
| 3 | 120.113,474 | 177.668,682 |
| 4 | 177.668,683 | 235.223,891 |
| 5 | 235.223,892 | 292.779,099 |

Hasil dari pembobotan diperoleh dataset baru sebagai berikut:

Tabel 4. Dataset Hasil Pembobotan

| No | KdCus | R | F | M |
|-----|----------|-----|-----|-----|
| 1 | BGL | 5 | 1 | 1 |
| 2 | BGL 1310 | 5 | 1 | 1 |
| 3 | BGL 965 | 5 | 5 | 2 |
| 4 | BGL001 | 4 | 1 | 2 |
| 5 | BGL0012 | 2 | 1 | 1 |
| 6 | BGL002 | 5 | 2 | 1 |
| 7 | BGL003 | 5 | 1 | 1 |
| 8 | BGL005 | 5 | 2 | 1 |
| 9 | BGL01 | 4 | 1 | 1 |
| 10 | BGL011 | 5 | 2 | 1 |
| 11 | BGL02 | 5 | 3 | 1 |
| 12 | BGL04 | 5 | 2 | 1 |
| 13 | BGL05 | 5 | 4 | 2 |
| 14 | BGL06 | 5 | 2 | 1 |
| 15 | BGL1 | 4 | 2 | 1 |
| 16 | BGL10 | 5 | 5 | 2 |
| 17 | BGL1005 | 4 | 1 | 1 |
| 18 | BGL101 | 5 | 1 | 1 |
| 19 | BGL1011 | 1 | 1 | 1 |
| 20 | BGL102 | 2 | 1 | 1 |
| ... | ... | ... | ... | ... |

3.2. Tahap Clustering

Tahap clustering adalah tahap mengelompokkan customer ke dalam 5 katagori loyalitas. Untuk *clustering* data ini

digunakan algoritma K-Mean dimana ditentukan K=5 sesuai katagori yang diinginkan dan teknik perhitungan jarak menggunakan *Euclidean Distance*. Dengan menggunakan algoritma K-Means setelah iterasi ke-6 setiap data tidak berpindah lagi ke cluster lain atau sama dengan hasil iterasi sebelumnya, berarti kondisi cluster sudah mencapai konvergen dan proses iterasi pun berhenti.

Tabel 5. Hasil Iterasi Terakhir (6)

| No | IdCus | Atribut | | | Ikut Cluster |
|-----|----------|---------|-----|-----|--------------|
| | | R | F | M | |
| 1 | BGL | 5 | 1 | 1 | 2 |
| 2 | BGL.1310 | 5 | 1 | 1 | 2 |
| 3 | BGL.965 | 5 | 5 | 2 | 4 |
| 4 | BGL001 | 4 | 1 | 2 | 2 |
| 5 | BGL0012 | 2 | 1 | 1 | 1 |
| 6 | BGL002 | 5 | 2 | 1 | 3 |
| 7 | BGL003 | 5 | 1 | 1 | 2 |
| 8 | BGL005 | 5 | 2 | 1 | 3 |
| 9 | BGL01 | 4 | 1 | 1 | 2 |
| 10 | BGL011 | 5 | 2 | 1 | 3 |
| 11 | BGL02 | 5 | 3 | 1 | 3 |
| 12 | BGL04 | 5 | 2 | 1 | 3 |
| 13 | BGL05 | 5 | 4 | 2 | 4 |
| 14 | BGL06 | 5 | 2 | 1 | 3 |
| 15 | BGL1 | 4 | 2 | 1 | 2 |
| 16 | BGL10 | 5 | 5 | 2 | 4 |
| 17 | BGL1005 | 4 | 1 | 1 | 2 |
| 18 | BGL101 | 5 | 1 | 1 | 2 |
| 19 | BGL1011 | 1 | 1 | 1 | 1 |
| 20 | BGL102 | 2 | 1 | 1 | 1 |
| ... | ... | ... | ... | ... | ... |

Tabel 6. Centroid akhir yang diperoleh

| Centroid | X ₁ | X ₂ | X ₃ |
|----------|----------------|----------------|----------------|
| 1 | 2.14706 | 1.03125 | 3.77273 |
| 2 | 1.17647 | 5.00000 | 1.56818 |
| 3 | 1.02941 | 2.28311 | 5.00000 |
| 4 | 4.52083 | 1.05023 | 2.42857 |
| 5 | 1.15625 | 4.79545 | 3.78571 |

Tabel 7. Keanggotaan setiap Cluster

| Keanggotaan Tiap Cluster | | |
|--------------------------|--------------|--------|
| Cluster | Status | Jumlah |
| 1 | Tidak Loyal | 34 |
| 2 | Kurang Loyal | 192 |
| 3 | Biasa | 219 |
| 4 | Loyal | 44 |
| 5 | Sangat Loyal | 14 |
| Total | | 503 |

3.3. Tahap Evaluation

Untuk evaluasi DBI diawali dengan menghitung SSW. SSW (*Sum of Square Within Cluster*) adalah cara menghitung jarak *intra-cluster* sebagai metrik kohesi dalam sebuah cluster ke-*i* dengan menghitung rata-rata jarak *euclidean* tiap data terhadap cluster ke-*i*. Untuk menghitung SSW data harus dikelompokkan berdasarkan *cluster* yang dihasilkan. Dengan perhitungan SSW:

$$SSW_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_j, c_i)$$

Dimana m_i adalah jumlah data yang berada dalam cluster ke-*i*, sedangkan C_i adalah centroid cluster ke-*i*. Maka diperoleh hasil perhitungan SSW:

Tabel 8. Hasil Perhitungan SSW

| Perhitungan SSW | | | |
|-----------------|-----------|---------|---------|
| Centroid | Jumlah | | SSW |
| | Jarak | Anggota | |
| 1 | 24.58501 | 34 | 0.72309 |
| 2 | 117.29435 | 192 | 0.61091 |
| 3 | 97.43748 | 219 | 0.44492 |
| 4 | 39.47473 | 44 | 0.89715 |
| 5 | 17.43099 | 14 | 1.24507 |

Sedang untuk menghitung separasi antar *cluster* atau jarak *inter-cluster* misalnya *cluster i* dan *j* digunakan formula *Sum of Square Between Cluster (SSB)* dengan mengukur jarak antar centroid c_i dan c_j seperti pada persamaan (2.4):

$$SSB_{i,j} = d(c_i, c_j)$$

SSB diperoleh dengan menghitung jarak antar-centroid. Maka diperoleh hasil perhitungan SSB:

Tabel 9. Hasil Perhitungan SSB

| | | Perhitungan SSB | | | | |
|---------|---------|-----------------|---------|---------|---------|---|
| | | Cluster | | | | |
| Cluster | SSB | 1 | 2 | 3 | 4 | 5 |
| 1 | 0.00000 | 2.37386 | 3.06012 | 3.74764 | 4.15981 | |
| 2 | 2.37386 | 0.00000 | 1.23465 | 2.68508 | 3.07172 | |
| 3 | 3.06012 | 1.22465 | 0.00000 | 1.59031 | 2.73924 | |
| 4 | 3.74764 | 2.68508 | 1.59031 | 0.00000 | 2.60116 | |
| 5 | 4.15983 | 3.07172 | 2.73924 | 2.60116 | 0.00000 | |

Sementara nilai DBI didapatkan dari persamaan berikut berikut:

$$DBI = \frac{1}{K} \sum_{i=1}^k \max_{i \neq j} (R_{ij})$$

dengan hasil DBI sebagai berikut:

| | | Perhitungan R dan DBI (Berdasarkan Indeks) | | | | | |
|---|---------|--|---------|---------|---------|---------|---------|
| | | Cluster | | | | | |
| R | 1 | 2 | 3 | 4 | 5 | R Max | DBI |
| 1 | 0.00000 | 0.56190 | 0.38189 | 0.43234 | 0.47313 | 0.56190 | 0.79074 |
| 2 | 0.56190 | 0.00000 | 0.38214 | 0.50364 | 0.46422 | 0.50364 | |
| 3 | 0.38189 | 0.38214 | 0.00000 | 0.44391 | 0.49693 | 0.49693 | |
| 4 | 0.43234 | 0.50364 | 0.44391 | 0.00000 | 0.51196 | 0.51196 | |
| 5 | 0.47313 | 0.46422 | 0.49693 | 0.51196 | 0.00000 | 0.51196 | |

Dari hasil perhitungan diperoleh nilai DBI sebesar 0.79074 (Non-Negatif)

4. SIMPULAN

Berdasarkan tabel 10 hasil DBI sebesar 0.79074 menggambarkan clustering terhadap loyalitas customer terbentuk cukup baik. Karena secara esensial DBI menginginkan nilai sekecil mungkin (non-negatif=0) untuk menilai baiknya cluster yang didapat. Nilai tersebut mengindikasikan kualitas keanggotaan dalam sebuah cluster (*intra-cluster*) memiliki tingkat kemiripan yang cukup tinggi dan jarak ketidakmiripan antar cluster (*inter-luster*) yang juga cukup tinggi.

Komposisi katagori loyalitas customer sebagaimana tabel 7 menggambarkan bahwa dengan metode RFM dapat diketahui masih rendahnya nilai loyalitas customer. Untuk itu diperlukan strategi manajemen yang lebih baik dalam meningkatkan loyalitas customernya.

5. DAFTAR PUSTAKA

- Widiarina dan R. S. Wahono, "Algoritma Cluster Dinamik untuk Optimasi Cluster pada Algoritma K-Means dalam Pemetaan Nasabah Potensial," *Journal of Intelligent Systems*, 2015.
- R. Hallowell, "The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study," *International Journal of Service Industry Management*, 1996.
- T. Musanto, "Faktor-Faktor Kepuasan Pelanggan dan Loyalitas Pelanggan: Studi Kasus pada CV. Sarana Media Advertising Surabaya," *Jurnal Manajemen & Kewirausahaan*, 2004.
- H. Yusuf, *Dasar-dasar Akuntansi*, Yogyakarta: Bagian Penerbitan Sekolah Tinggi Ilmu Ekonomi UGM, 2001.
- Mulyadi, *Sistem Akuntansi*, Cetakan ketiga, Agustus 2001, UGM Yogyakarta: Salemba Empat (PT Salemba Emban Patria), 2001.
- M. Chandrashekar, "Satisfaction Strength and Customer Loyalty," *Journal of Marketing Research Vol. XLIV (February 2007)*, 2007.
- E. Ngai dan D. C. Li Xiu, "Application of data mining techniques in customer relationship management: A literature review and classification," *Expert Systems with Applications*, 2009.
- C. H. Cheng dan Y. S. Chen, "Classifying the segmentation of customer value via RFM model and

Field Code Changed

- RS theory," *Expert Systems with Applications*, 2009.
- [9] W. Jo-Ting, L. Shih-Yen dan W. Hsin-Hung. "A review of the application of RFM model," *African Journal of Business Management*, 2010.
- [10] Kaymak, "Fuzzy target selection using RFM variables," *IFSA World Congress and 20th NAFIPS International Conference*, 2001.
- [11] S. M. Hosseini dan A. Maleki, "Cluster analysis using data mining approach to develop CRM methodology," *Expert Systems with Applications*, 2010.
- [12] E. Prasetyo, *Data Mining. Mengolah Data Menjadi Informasi Menggunakan Matlab*, Yogyakarta: Andi Offset Yogyakarta, 2014.
- [13] B. Santoso, *Data Mining - Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Graha Ilmu Cetakan I, 2007.
- [14] F. A. Hermawati, *Data Mining*, Yogyakarta: Andi Offset, 2013.
- [15] U. Fayyad, "From Data Mining to Knowledge Discovery in Databases," *AI Magazine Volume 17 Number 3*, 1996.

ORIGINALITY REPORT

13%

SIMILARITY INDEX

9%

INTERNET SOURCES

4%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES

| | | |
|---|---|-----|
| 1 | docobook.com Internet Source | 1% |
| 2 | Submitted to UIN Sultan Syarif Kasim Riau Student Paper | 1% |
| 3 | www.jp.feb.unsoed.ac.id Internet Source | 1% |
| 4 | repository.ipb.ac.id Internet Source | 1% |
| 5 | www.scribd.com Internet Source | 1% |
| 6 | eprints.dinus.ac.id Internet Source | 1% |
| 7 | Submitted to Universitas Dian Nuswantoro Student Paper | 1% |
| 8 | Ni Wayan Wardani, Gede Rasben Dantes, Gede Indrawan. "Prediksi Customer Churn dengan Algoritma Decision Tree C4.5 Berdasarkan Segmentasi Pelanggan untuk | <1% |

Mempertahankan Pelanggan pada Perusahaan Retail", Jurnal RESISTOR (Rekayasa Sistem Komputer), 2018

Publication

| | | |
|----|---|-----|
| 9 | Submitted to St. John Fisher College Student Paper | <1% |
| 10 | www.ijimai.org Internet Source | <1% |
| 11 | edoc.site Internet Source | <1% |
| 12 | amiktaranaprobolinggo.blogspot.com Internet Source | <1% |
| 13 | libetd.shu.edu.tw Internet Source | <1% |
| 14 | tr.scribd.com Internet Source | <1% |
| 15 | www.contohjurnal.net Internet Source | <1% |
| 16 | id.scribd.com Internet Source | <1% |
| 17 | docplayer.info Internet Source | <1% |
| 18 | Submitted to City University of Hong Kong Student Paper | <1% |

19

Submitted to Padjadjaran University

Student Paper

<1%

20

Submitted to Syiah Kuala University

Student Paper

<1%

21

Submitted to Universitas Muria Kudus

Student Paper

<1%

22

Muhammet Pakyurek, Mehmet Selman Sezgin, Sedat Kestepc, Busra Bora, Remzi Duzagac, Olcay Taner Yildiz. "Customer clustering using RFM analysis", 2018 26th Signal Processing and Communications Applications Conference (SIU), 2018

Publication

<1%

23

www.thinkmind.org

Internet Source

<1%

24

zadoco.site

Internet Source

<1%

25

bitkom-ti.blogspot.com

Internet Source

<1%

26

Submitted to Universitas Pelita Harapan

Student Paper

<1%

27

www.iraj.in

Internet Source

<1%

28

Submitted to Universitas Atma Jaya Yogyakarta

Student Paper

<1%

29

Submitted to Universitas Negeri Jakarta

Student Paper

<1%

30

Ramón A. Carrasco, María Francisca Blasco, Enrique Herrera-Viedma. "A 2-tuple Fuzzy Linguistic RFM Model and Its Implementation", Procedia Computer Science, 2015

Publication

<1%

31

banik.ir

Internet Source

<1%

32

media.neliti.com

Internet Source

<1%

Exclude quotes On

Exclude matches Off

Exclude bibliography Off

IEEE Journal

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9
