

## SEGMENTASI PELANGGAN RITEL PRODUK FARMASI OBAT MENGGUNAKAN METODE DATA MINING KLASTERISASI DENGAN ANALISIS REGENCY FREQUENCY MONETARY (RFM) TERMODIFIKASI

Arief Wibowo\*<sup>1</sup>, Andy Rio Handoko<sup>2</sup>

<sup>1,2</sup>Fakultas Teknologi Informasi - Universitas Budi Luhur  
Email: <sup>1</sup>arief.wibowo@budiluhur.ac.id, <sup>2</sup>andy.handoko@budiluhur.ac.id  
\*Penulis Korespondensi

(Naskah masuk: 21 Desember 2019, diterima untuk diterbitkan: 27 April 2020)

### Abstrak

Secara umum, pembelian produk farmasi di Indonesia tidak memiliki pola. Pembelian produk farmasi seperti obat-obatan, dilakukan oleh individu bukan sebagai persiapan untuk menjaga kesehatan, namun sebagai respon terhadap penyakit yang sedang diderita. Di sisi lain, pelanggan retail produk farmasi obat biasanya dipengaruhi oleh faktor harga jual dan faktor kecocokan (sugesti) pada merk obat tertentu sewaktu melakukan pembelian. Berdasarkan kondisi itu maka pola pembelian obat bagi masyarakat Indonesia menjadi tidak dapat diprediksi. Hal tersebut membuat pelaku usaha di bisnis ritel produk farmasi obat, relatif sulit untuk meningkatkan nilai penjualan. Salah satu upaya yang bisa dilakukan pelaku bisnis untuk meningkatkan pendapatan adalah dengan melakukan promosi penjualan berdasarkan jenis kelompok pelanggannya. Transaksi pembelian produk farmasi obat dapat dianalisis untuk mengetahui segmentasi pelanggan berdasarkan pola pembelian. Riset ini telah berhasil memodelkan segmentasi pelanggan ritel apotek dengan teknik data mining klasterisasi. Metode yang digunakan adalah melakukan analisis data transaksi pembelian yang terdiri dari atribut *Recency Frequency Monetary (RFM)* termodifikasi. Analisis telah melibatkan atribut Kuantitas (*Quantity*) dari data transaksi pembelian produk farmasi obat sebagai eksperimen modifikasi model. Pada proses pemodelan klasterisasi, studi ini menggunakan algoritme data mining K-Means. Hasil penelitian menunjukkan bahwa segmentasi pelanggan yang optimal berada pada dua klaster berdasarkan hasil analisis *QRF (Quantity, Recency dan Frequency)* menggunakan evaluasi *Davies Bouldin Indeks (DBI)* dengan nilai 0,527. Kinerja model tersebut dibandingkan dengan algoritme *K-Medoids*. Hasil klasterisasi pelanggan pada dua kategori menggunakan K-Medoids memiliki nilai DBI sebesar 1,334. Berdasarkan nilai perbandingan tersebut maka metode K-Means terbukti lebih baik dalam pembentukan klaster pelanggan ritel farmasi obat pada analisis atribut *Quantity, Recency dan Frequency*.

**Kata kunci:** *Segmentasi Pelanggan, Ritel Apotek, Analisis Recency Frequency Monetary (RFM) Termodifikasi, Data Mining Klasterisasi*

## SEGMENTATION OF CUSTOMERS OF DRUG PHARMACEUTICAL PRODUCT RETAIL USING CLUSTERIZATION MINING DATA METHOD USING MODIFIED MONETARY REGENCY FREQUENCY (RFM) ANALYSIS

### Abstract

In general, the purchase of pharmaceutical products in Indonesia has no pattern. The purchase of pharmaceutical products such as medicines, made by individuals not as preparation for maintaining health, but in response to the illness being suffered. On the other hand, retail customers of pharmaceutical drug products are usually influenced by selling price factors and suggestions for certain drug brands when making a purchase. Based on these conditions, the pattern of purchasing drugs for Indonesian people is unpredictable. This makes businesses in the retail business of pharmaceutical drug products, relatively difficult to increase sales value. One effort that businesses can do to increase revenue is to conduct sales promotions based on the type of customer group. Drug pharmaceutical product purchase transactions can be analyzed to determine customer segmentation based on purchase patterns. This research has successfully modeled the pharmacy retail customer segmentation with clustering data mining techniques. The method used is to analyze the purchase transaction data consisting of modified *Recency Frequency Monetary (RFM)* attributes. Analysis has involved the *Quantity* attribute (*Quantity*) of the transaction data of pharmaceutical drug product purchases as a model modification experiment. In the cluster modeling process, this study uses the *K-Means* data mining algorithm. The results showed that the optimal customer segmentation was in two clusters based on the results of the *QRF (Quantity, Recency and Frequency)*

*analysis using the Davies Bouldin Index (DBI) evaluation with a value of 0.527. The performance of the model is compared with the K-Medoids algorithm. The results of customer clustering in two categories using K-Medoids have a DBI value of 1.334. Based on these comparative values, the K-Means method is proven to be better in forming pharmaceutical drug retail customer clusters with analysis Quantity, Recency and Frequency attributes.*

**Keywords:** *Customer Segmentation, Pharmacy Retail, Modified of Recency Frequency Monetary (RFM) Analysis, Data Mining Clustering*

## 1. PENDAHULUAN

Meningkatnya kesadaran masyarakat Indonesia tentang hidup yang sehat membawa perubahan pada pola kunjungan ke fasilitas kesehatan yang tersedia. Hadirnya institusi yang memberikan jaminan kesehatan masyarakat Indonesia (BPJS) sejak tahun 2014 ternyata belum mampu memenuhi kebutuhan layanan kesehatan disertai keterbatasan obat yang diberikan. Di sisi lain, sebagian besar masyarakat Indonesia masih melakukan Swamedikasi yaitu upaya pengobatan yang dilakukan diri sendiri tanpa melalui resep dokter (Jajuli dan Sinuraya, 2018). Kondisi-kondisi tersebut turut membentuk pola masyarakat dalam pembelian obat maupun produk farmasi di apotek.

Faktor-faktor yang memengaruhi pembelian obat di masyarakat, antara lain pengaruh iklan atau informasi untuk mendapatkan obat dengan biaya murah dan cepat tanpa resep dokter (Jajuli dan Sinuraya, 2018), tingkat pengetahuan tentang sakit yang diderita (Restiyono, 2016) atau pengetahuan tentang obat yang akan dibeli (Syafitri, Hidayati dan Pristianty, 2017). Dengan kondisi-kondisi tersebut, maka referensi pembelian obat dari tenaga medis di fasilitas layanan kesehatan, bukan satu-satunya faktor pendorong pembentukan pola pembelian obat dan hal ini menjadi tantangan bagi pebisnis apotek dalam mempertahankan kelangsungan usahanya.

Dengan berbagai faktor penentu pada keputusan pembelian obat, maka peritel obat seperti apotek harus memiliki strategi khusus dalam peningkatan nilai penjualan produk farmasi di apotek. Studi-studi terdahulu telah dilakukan untuk memecahkan persoalan tersebut, misalnya strategi bauran pemasaran (Dysyandi et al., 2019), analisis asosiasi pembelian produk farmasi (Febrian et al., 2018), atau analisis pembentukan segmentasi atau klasterisasi pelanggan produk ritel farmasi (Hendrawan, Utamima dan Husna, 2015).

Proses klasterisasi pada studi Hendrawan pada tahun 2015, dilakukan untuk mengetahui konsumen atau pelanggan yang memiliki nilai tertentu yang dapat dijadikan sasaran promosi atau pemasaran produk, dengan analisis RFM (*Recency, Frequency, Monetary*) dan *Rough-set*. Menarik untuk diketahui apakah variabel jumlah pembelian (*quantity*) turut mempengaruhi pembentukan klaster pelanggan apotek, karena faktor pembelian obat sangat beragam, dan tidak terpola sebagaimana studi terdahulu yang telah disampaikan.

Dilatarbelakangi oleh masih minimnya jumlah penelitian tentang segmentasi pelanggan apotek yang menggunakan analisis RFM, maka studi ini akan mengusulkan model baru klasterisasi pelanggan berdasarkan data transaksi untuk variabel *Recency, Frequency* dan *Monetary* (RFM) termodifikasi. Modifikasi pada variabel model dilakukan dengan memasukkan atribut *Quantity* untuk mendapatkan hasil klasterisasi yang lebih baik. Evaluasi model klasterisasi yang terbentuk pada studi ini akan diuji menggunakan *Davies Bouldin Index* dan uji kinerja algoritma pembandingan menggunakan metode *K-Medoids*. Proses segmentasi pelanggan yang lebih baik akan menghasilkan rekomendasi strategi pemasaran yang lebih efektif untuk meningkatkan nilai penjualan.

## 2. STUDI LITERATUR

Pelanggan termasuk salah satu aset yang harus dikelola oleh pengusaha dalam mempertahankan bisnisnya. Perusahaan harus menjaga pelanggan yang dimiliki untuk dapat terus berjalan dan berkembang. Semakin banyak dan loyal pelanggan yang dimiliki oleh perusahaan, maka semakin baik untuk perusahaan. Namun sifat pelanggan yang selalu pilih-pilih membuat perusahaan kesulitan dalam membedakan pelanggan yang dapat memberikan keuntungan besar bagi perusahaan dan yang tidak atau kurang menguntungkan bagi perusahaan (Hidayatullah, Rokhmawati dan Perdanakusuma, 2018). Segmentasi pelanggan dalam buku yang ditulis oleh Don Pepper & Martha Roger berjudul *Managing Customer Relationship: A Strategic Framework*, terdiri dari:

- a. *Most Valuable Customer (MVC)*, merupakan jenis pelanggan yang bernilai paling tinggi bagi perusahaan. Jenis pelanggan pada kategori ini yang memberikan keuntungan paling banyak bagi perusahaan. Pelanggan jenis ini telah diketahui dari transaksinya.
- b. *Most Growable Customer*, merupakan pelanggan yang memiliki potensi besar, namun belum disadari hal tersebut merupakan keuntungan bagi perusahaan karena perlu dianalisis lebih lanjut.
- c. *Below Zeros*, merupakan jenis pelanggan yang memberikan keuntungan lebih sedikit dari biaya yang dikeluarkan untuk memberikan pelayanan.
- d. *Migrators*, merupakan pelanggan yang berada pada posisi antara *below zeros* dan *most growable customer*.

Dengan mengetahui klusterisasi pelanggan maka akan dapat direncanakan usaha-usaha peningkatan pendapatan berdasarkan promosi yang sesuai dari kelompok pelanggan. Salah satu teknik klusterisasi yang dapat dilakukan adalah teknik penambangan data.

Data mining adalah serangkaian proses untuk menemukan pola yang menarik berupa aturan atau pengetahuan dari data yang jumlahnya besar (Ramadhan, Efendi dan Mustakim, 2017). Data mining termasuk ke dalam disiplin ilmu yang mempelajari metode ekstraksi pengetahuan atau penemuan pola dari suatu data (Butarbutar dkk., 2016). Salah satu pekerjaan dalam penambangan data adalah klusterisasi. Klusterisasi merupakan kegiatan mengelompokkan data, hasil observasi dan kasus dalam kelas yang memiliki kemiripan (Butarbutar dkk., 2016).

Algoritme K-Means yang ditemukan oleh MacQueen pada tahun 1967 merupakan algoritme pengelompokan iteratif yang bekerja dengan cara melakukan partisi dataset ke sejumlah k-kluster yang ditetapkan sebelumnya (Hidayatullah, Rokhmawati dan Perdanakusuma, 2018). Data dengan tingkat kemiripan yang tinggi maka akan dikelompokkan ke dalam sebuah kluster, dan data yang memiliki kemiripan berbeda dikelompokkan ke dalam kluster lainnya yang berbeda. *Partitioning* pada algoritme K-Means digunakan terlebih dahulu untuk menentukan jumlah kluster akhir (Yohanes, 2016).

Algoritme K-Medoids bekerja dengan cara memakai perwakilan untuk mengatasi kelemahan dari algoritme K-Means. Algoritma ini menghilangkan pemakaian rata-rata dalam pembaharuan centroid dan menggantinya dengan objek aktual sebagai representasi suatu kluster. K-Medoids melakukan partisi dengan cara meminimalkan jumlah dissimilarity antara objek  $p$  dan objek representatif terdekat, dengan menggunakan jumlah kesalahan absolut. Pelibatan variabel  $E$  merupakan jumlah kesalahan absolut untuk semua objek  $p$  pada suatu himpunan data, dan  $o_i$  merupakan objek representatif dari kluster  $C_i$ .

Konsekuensi yang muncul jika menggunakan cara ini adalah kompleksitas komputasi menjadi lebih tinggi karena harus melakukan pengecekan penggantian objek representatif dengan setiap objek lain yang bukan representatif akan meningkatkan kualitas klusterisasi. Pada umumnya K-Medoids diimplementasikan menggunakan algoritma PAM (*Partitioning Around Medoids*) yang mirip dengan algoritma K-Means (Suyanto, 2017).

Pada studi terdahulu (Luo et al., 2012) telah dilakukan penelitian mengenai klusterisasi pelanggan untuk Changzhou Telecom di Provinsi Jiangsu yang melakukan analisis data pelanggan. Hasil klusterisasi digunakan dalam strategi pemasaran berdasarkan pemodelan dengan algoritme K-Means. Batasan penelitian ini adalah tidak dilakukannya validasi pada hasil klusterisasi, namun demikian peneliti

menyatakan bahwa segmentasi yang dilakukan telah diimplementasikan pada obyek studi.

Penelitian lain dilakukan mengenai penentuan strategi pemasaran produk menggunakan algoritme K-Medoids (Triyanto, 2015) yang menghasilkan lima buah kluster pada dataset penjualan. Studi ini memiliki batasan, berupa tidak dilakukannya evaluasi kluster yang terbentuk berdasarkan analisis RFM.

Peneliti lain menggunakan algoritme Kluster Dinamik untuk mengoptimasi kluster pada metode K-Means dalam pemetaan nasabah yang potensial (Widiarina dan Romi Satria Wahono, 2015). Studi tersebut dilakukan untuk pembentukan kluster yang optimal menggunakan metode RFM pada proses persiapan data. Penelitian ini menggunakan *Davies Bouldin Index* sebagai pengukuran validitas kluster. Penelitian ini cukup baik karena mengoptimasi proses klusterisasi menggunakan analisis RFM, dan belum melibatkan atribut lain yang berkemungkinan meningkatkan optimalisasi kluster yang terbentuk.

Model analisis RFM sendiri dikembangkan pertama kali oleh Hughes (Hughes, 1994) sebagai metode untuk menganalisis nilai pelanggan untuk membentuk segmentasi pelanggan. Segmentasi pelanggan dibentuk berdasarkan data pelanggan menggunakan tiga atribut yaitu selang waktu (interval) pemakaian pelanggan, frekuensi dan jumlah uang (Hendrawan, Utamima dan Husna, 2015). Secara detail ketiga atribut tersebut dideskripsikan seperti di bawah ini:

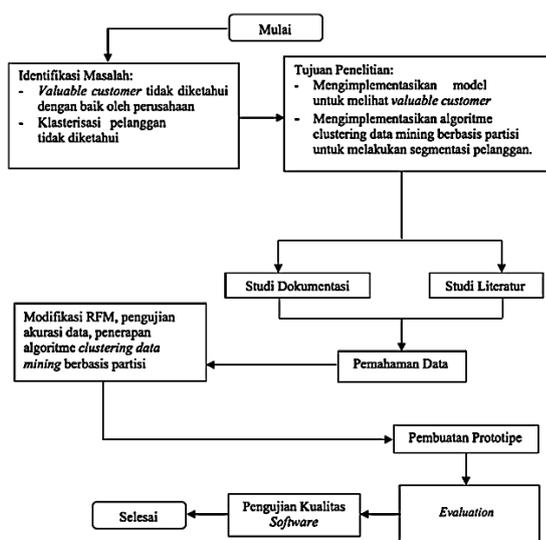
- a. *Recency of the last Purchase* (R). Atribut R merepresentasikan resensi, yaitu jarak antara waktu terakhir pemakaian/pembelian dengan waktu sekarang.
- b. *Frequency of the purchases* (F). Atribut F merepresentasikan frekuensi, yaitu jumlah transaksi pada periode waktu tertentu.
- c. *Monetary value of the purchases* (M). Atribut M merepresentasikan moneter, yaitu jumlah uang yang dibelanjakan pelanggan untuk melakukan pembelian pada periode waktu tertentu.

*Davies-Bouldin Index* (DBI) adalah salah satu metode validasi kluster dari hasil pemodelan klusterisasi. Pendekatan pengukuran DBI bekerja dengan cara memaksimalkan jarak antar kluster dan meminimalkan jarak intra kluster. Semakin kecil nilai DBI (positif mendekati nol) menunjukkan skema kluster yang lebih optimal (Widiarina dan Romi Satria Wahono, 2015).

Mereferensi pada studi terdahulu, maka perbedaan studi ini terdapat pada proses segmentasi pelanggan yang menerapkan metode RFM termodifikasi dengan pelibatan atribut *Quantity* (jumlah pembelian) untuk mendapatkan luaran proses klusterisasi. Dengan pelibatan atribut *Quantity* selain RFM, diduga akan membentuk kluster yang lebih optimum. Evaluasi model klusterisasi dilakukan pada studi ini, menggunakan *Davies-Bouldin Index* mengacu pada penelitian terdahulu.

### 3. METODOLOGI PENELITIAN

Data yang akan dianalisis pada studi ini berasal dari salah satu bisnis ritel yang menjual obat berdasarkan resep dokter, obat bebas, vitamin, suplemen, perlengkapan dan alat kesehatan, tes kesehatan, konsultasi Apoteker dan praktek dokter umum. Data yang digunakan adalah data transaksi penjualan vitamin di suatu Apotek berlokasi di Jakarta selama 12 bulan lebih dari September 2018 hingga November 2019. Seluruh tahap penelitian terlihat pada Gambar 1.



Gambar 1. Tahap-Tahap Penelitian

Pada Gambar 1 terlihat bahwa studi ini akan memodifikasi variabel model RFM dan melakukan proses klusterisasi dengan algoritma data mining. Proses pemodelan akan dilakukan berulang kali dengan menguji coba kombinasi variabel model RFM dengan *Quantity*, untuk mendapatkan hasil klusterisasi yang terbaik.

### 4. HASIL DAN PEMBAHASAN

Berdasarkan penyelesaian seluruh tahapan penelitian, maka didapatkan hasil sebagaimana yang diharapkan, yaitu terbentuknya klusterisasi pelanggan yang optimal, diketahui dari nilai DBI yang dijadikan sebagai standar evaluasi untuk kluster yang terbentuk.

#### 4.1. Pra-pemrosesan data

Jenis data transaksi yang dimiliki Apotek yang menjadi obyek penelitian pada studi kali ini adalah data transaksi dengan satu atau banyak record dalam satu nomor transaksi, dan jumlah produk dalam satu record terdapat satu atau lebih. Pembeli yang melakukan transaksi sebanyak 82% adalah member dan 18% non member.

Data sekunder memiliki delapan atribut yang kemudian dilakukan seleksi pemilihan atribut sesuai dengan tujuan dari penelitian ini, yaitu untuk menganalisis faktor RFM dan melibatkan atribut lain,

dengan dataset pelatihan yang terdiri dari atribut *id\_transaksi*, *kode\_pelanggan*, *tanggal\_transaksi*, *kuantitas* dan *harga*. Dari keseluruhan atribut tersebut dilakukan eksperimen pemodelan klusterisasi menggunakan atribut *Recency Frequency Monetary* dan *Quantity*.

#### 4.2. Penyiapan Data

Data sekunder terdiri dari 2.030 record data transaksi dengan 1.013 pelanggan yang terjadi selama 61 hari terhitung pada periode 1 September 2018 sampai dengan 31 November 2019. Penyiapan data dilakukan dengan melakukan pengolahan nilai dari setiap atribut untuk mendapatkan nilai QRFM (*Quantity, Recency, Frequency, Monetary*).

Setelah dilakukan seleksi atribut maka dilakukan pembersihan data dengan menghilangkan *missing value* yang disebabkan oleh transaksi yang dilakukan oleh pembeli non member. Hal ini dilakukan untuk mendapatkan data terpilih yang dapat dinyatakan pelanggan loyal.

Pada tahap selanjutnya proses transformasi data dilakukan dengan melakukan perhitungan total kuantitas produk yang dibeli, dan frekuensi transaksi berapa banyak tanggal transaksi yang berbeda pada setiap pelanggan dalam satu periode data yang telah ditentukan. Untuk mengetahui nilai *recency* dilakukan pencarian tanggal terakhir transaksi pembelian dalam satu periode yang ditentukan dan dibandingkan dengan tanggal saat analisis data dilakukan. Hasil penyiapan data untuk tahap pemodelan terlihat pada Tabel 1.

Tabel 1. Dataset primer untuk tahap pemodelan

<i>Cust_id</i>	<i>Qty</i>	<i>Last_order</i>	<i>Trx_count</i>	<i>Amount</i>
HRK-0102	16	13/06/18	11	226
HRK-0103	8	11/03/18	5	207
HRK-0104	3	03/06/19	2	195
HRK-0105	27	15/04/18	18	252
HRK-0106	24	04/07/18	16	245
HRK-0107	33	15/07/18	22	266
HRK-0108	89	26/06/19	59	1.011
HRK-0109	22	01/10/19	15	240
HRK-0110	28	17/11/19	19	254
HRK-0111	46	14/10/19	31	297
HRK-0112	13	29/07/19	9	219
HRK-0113	6	31/08/19	4	202
HRK-0114	81	05/09/19	54	956

Dari Tabel 1 terlihat bahwa ada enam atribut pada data primer untuk pemodelan klusterisasi, meliputi *Cust\_id*, *Qty*, *Last\_order*, *Trx\_count* dan *Amount*. Untuk atribut *Amount* telah dilakukan transformasi dengan melakukan *scaling* berbasis nilai ribuan rupiah dari data sekunder.

#### 4.3. Pemodelan Klusterisasi

Pada tahapan pemodelan digunakan algoritma data mining klusterisasi berbasis partisi. Pembentukan model dilakukan dengan bantuan perangkat lunak Rapidminer untuk mengetahui nilai DBI yang terbaik di antara dua algoritma K-Means

dan K-Medoids. Pemilihan didasarkan hasil uji dengan DBI, bahwa semakin rendah/kecil nilai indeks maka semakin optimal skema klaster yang terbentuk. Eksperimen pemodelan akan menggunakan pembagian dua klaster, sesuai dengan tujuan penelitian yang ditetapkan berdasarkan kebutuhan pengguna (pihak manajemen) dari Apotek.

#### 4.3.1. Pemodelan dengan atribut RFM

Pada proses pemodelan kedua dengan menggunakan atribut RFM, menghasilkan perubahan pada keanggotaan klaster dan nilai DBI yang dapat dilihat pada Tabel 2.

Tabel 2. Pemodelan Menggunakan Atribut RFM

Jumlah k-	Nilai DBI
2	0,529
3	0,802
4	0,626

Dari Tabel 2 terlihat bahwa nilai DBI paling optimal yang dihasilkan saat dilakukan pemodelan menggunakan atribut RFM berada pada jumlah klaster sebanyak 2 dengan nilai DBI sebesar 0,529. Hingga tahap ini, nilai DBI yang dihasilkan lebih optimal jika dibandingkan dengan klasterisasi menggunakan atribut QRFM.

#### 4.3.2. Pemodelan dengan atribut QRFM

Proses pemodelan data sekunder yang telah melalui proses data preparation, untuk mendapatkan klaster yang optimum berdasarkan nilai Davies Bouldin Index yang dihasilkan dengan bantuan framework data mining Rapidminer. Pada percobaan pertama menggunakan atribut QRFM. Nilai Davies Bouldin Index (DBI) yang dihasilkan jika menggunakan atribut QRFM terdapat dalam Tabel 3.

Tabel 3. Pemodelan Menggunakan Atribut QRFM

Jumlah k-	Nilai DBI
2	0,539
3	0,903
4	$\infty$

Dari Tabel 3 diketahui bahwa proses klasterisasi menggunakan atribut QRFM yang paling optimal berada pada jumlah klaster sebanyak 2 (dua) yang ditunjukkan dengan DBI sebesar 0,539.

#### 4.3.3. Pemodelan dengan atribut QFM

Pemodelan ketiga dilakukan dengan menggunakan tiga atribut yaitu QFM, menghasilkan perubahan pada keanggotaan klaster dan nilai DBI yang dapat dilihat pada Tabel 4.

Tabel 4. Pemodelan Menggunakan Atribut QFM

Jumlah k-	Nilai DBI
2	0,723
3	1,029
4	$\infty$

Nilai DBI paling optimal yang dihasilkan pada pemodelan menggunakan atribut QFM adalah dengan jumlah klaster sebanyak 2 dengan nilai DBI sebesar 0,729. Pada eksperimen ini, tidak ada nilai yang lebih baik daripada analisis dengan atribut QRFM maupun RFM.

#### 4.3.4. Pemodelan dengan atribut QRM

Pemodelan keempat dilakukan dengan menganalisis tiga atribut yaitu QRM, menghasilkan perubahan pada keanggotaan klaster dan nilai DBI yang dapat dilihat pada Tabel 4.

Tabel 4. Pemodelan Menggunakan Atribut QRM

Jumlah k-	Nilai DBI
2	0,532
3	0,806
4	1,046

Nilai DBI paling optimal yang dihasilkan pada pemodelan menggunakan atribut QFM adalah dengan jumlah klaster sebanyak 2 dengan nilai DBI sebesar 0,532. Pada eksperimen ini, model sedikit lebih baik dari eksperimen pada atribut QRFM dan QFM namun tidak lebih baik daripada analisis dengan atribut RFM.

#### 4.3.5. Pemodelan dengan atribut QRF

Pemodelan kelima dilakukan menggunakan tiga buah atribut yaitu QRF, menghasilkan perubahan pada keanggotaan klaster dan nilai DBI yang dapat dilihat pada Tabel 5.

Tabel 5. Pemodelan Menggunakan Atribut QRF

Jumlah k-	Nilai DBI
2	0,527
3	0,782
4	0,597

Nilai DBI paling optimum yang dihasilkan pada analisis atribut QRF adalah pada dua klaster dengan nilai DBI sebesar 0,527. Perbandingan seluruh hasil eksperimen pemodelan terbaik dua klaster yang telah dilakukan dapat dilihat pada Tabel 6.

Tabel 6. Perbandingan Hasil Pemodelan Terbaik Dua Klaster

Atribut	Nilai DBI
RFM	0,529
QRFM	0,539
QFM	0,723
QRM	0,532
QRF	0,527

Berdasarkan nilai DBI yang dihasilkan, ternyata klaster yang terbentuk paling optimum diantara hasil pemodelan dengan modifikasi atribut RFM, dengan menambahkan atribut Quantity, adalah pemodelan dengan analisis atribut QRF (*Quantity, Recency dan Frequency*) dengan nilai DBI sebesar 0,527.

#### 4.3.6. Pengujian dan Evaluasi

Studi ini juga melakukan perbandingan kinerja algoritme K-Means. Perbandingan kinerja dilakukan dengan menerapkan analisis data yang sama dengan algoritme data mining clustering berbasis partisi lainnya yaitu algoritme K-Medoids. Hasil perbandingan kinerja terlihat pada Tabel 7.

Tabel 7. Perbandingan Pemodelan Algoritme K-Medoids

Atribut	Jumlah k-	Nilai DBI
K-Means	2	0,527
K-Medoids	2	1,334

Pada hasil percobaan di Tabel 7 menunjukkan nilai *Davies Bouldin Index* yang dihasilkan algoritme K-Medoids adalah sebesar 1,334. Nilai DBI yang dihasilkan tersebut lebih tinggi dari nilai DBI pada percobaan menggunakan dua kluster dengan algoritme K-Means.

Setelah dilakukan pemodelan dengan menggunakan atribut RFM, QRFM, QFM, QRM dan QRF dalam algoritme K-Means dengan beberapa pilihan jumlah kluster, ternyata kluster terbaik berdasarkan nilai *Davies Bouldin Index* terkecil diperoleh dari pemodelan menggunakan atribut QRF (*Quantity Recency dan Frequency*) dengan jumlah kluster sebanyak dua, pada nilai DBI sebesar 0,527. Nilai tersebut dikatakan baik karena merupakan nilai terkecil yang diperoleh dari seluruh eksperimen. Ukuran model klusterisasi menggunakan nilai DBI yang semakin kecil atau semakin mendekati nilai nol dikatakan semakin optimal pada kluster yang dihasilkan (Widiarina dan Romi Satria Wahono, 2015).

Analisis kinerja juga dilakukan dengan menguji kluster optimal yang dihasilkan dari tahap pemodelan algoritme K-Means, dengan cara membandingkan dengan atribut QRF pada dua hingga empat kluster, menggunakan algoritme K-Medoids.

Dari hasil pemodelan yang telah dilakukan dapat diketahui bahwa nilai DBI pada algoritme K-Means ternyata lebih kecil atau lebih rendah dari pada menggunakan algoritme K-Medoids, yaitu sebesar 0,527 dengan kondisi optimal di dua kluster. Dengan demikian berdasarkan pemodelan yang dilakukan pada penelitian ini dapat disimpulkan bahwa algoritme K-Means merupakan algoritme yang lebih baik daripada K-Medoids dalam penentuan dua kluster pelanggan ritel di apotek ( $H_0$  diterima dan  $H_1$  ditolak).

Dari hasil klusterisasi K-Means dengan analisis QRF pada dua kluster, diketahui bahwa spesifikasi kluster pertama dari pelanggan memiliki indikator pembelian sebanyak tiga hingga tujuh item dalam satu kali transaksi pembelian (aspek *Quantity*), pembelian tidak pernah berjarak lebih dari 2 minggu dari pembelian sebelumnya (aspek *Recency*) dan melakukan *repeat order* setidaknya dua kali dalam satu bulan yaitu sebanyak tiga hingga lima kali pembelian ulang (aspek *Frequency*). Kategori

pelanggan di kluster pertama ini dapat disebut sebagai *Prime Customer*.

Kluster kedua dikategorikan *Regular Customer*, yaitu pelanggan yang hanya melakukan pembelian antara satu hingga dua item dalam satu kali transaksi pembelian, melakukan pembelian ulang paling banyak dua kali dengan rentang waktu satu hingga empat bulan periode pembelian dan *repeat order* dilakukan lebih dari dua bulan dari pembelian terakhir.

#### 4.3.6. Pengujian Prototipe

Pengujian kualitas aplikasi hasil implementasi model klusterisasi diselesaikan dengan metode *Black Box Testing* untuk mengetahui apakah aplikasi berjalan dengan baik sesuai fungsi dan kebutuhan pengguna.

Pengujian terhadap sistem, dilakukan dengan melakukan pengukuran kesesuaian sistem yang menjadi ekspektasi pengguna. Pendekatan *Software Quality Assurance* (SQA) yang digunakan memiliki beberapa tolak ukur, antara lain:

Tabel 8. Komponen Software Quality Assurance

No.	Metriks	Deskripsi
1.	<i>Auditability</i>	Kesesuaian pada standar
2.	<i>Completeness</i>	Kelengkapan kebutuhan sistem
3.	<i>Operability</i>	Kemudahan pengoperasian
4.	<i>Simplicity</i>	Kemudahan untuk dipahami
5.	<i>Training</i>	Kemudahan pembelajaran
6.	<i>Company Benefit</i>	Benefit bagi perusahaan
7.	<i>Time Management</i>	Efisiensi dalam penggunaan aplikasi

Dari tujuh komponen metrik pada Tabel 8, maka dibuat pertanyaan yang diberikan kepada *user* dalam kapasitas sebagai PiC (*Personal incharge Charge*) pada divisi penjualan di Apotek yang ditunjuk sebagai responden. Hasil evaluasi prototipe perangkat lunak yang dibangun beserta standar kriteria hasil pengujian terlihat pada Tabel 9 dan Tabel 10.

Tabel 9. Hasil Evaluasi Prototipe dengan SQA

No.	Responden	Skor Metrik Penilaian
1.	PiC 01	81,12%
2.	PiC 02	86,46%
3.	PiC 03	83,91%
4.	PiC 04	80,00%
5.	PiC 05	91,12%
6.	PiC 06	84,29%
7.	PiC 07	88,57%
8.	PiC 08	87,14%
9.	PiC 09	80,00%
10.	PiC 10	87,14%
	Rata-rata	84,29%

Tabel 10. Kriteria Hasil Pengukuran SQA

No.	% Jumlah Skor	Kriteria
1	0%-20%	Buruk
2.	20,01% - 36,00%	Tidak Baik
3.	36,01% - 52,00%	Kurang Baik
4.	52,01% - 68,00%	Cukup Baik
5.	68,01% - 84,00%	Baik
6.	84,01% - 100%	Sangat Baik

Dengan nilai prosentase rata-rata penilaian sebesar 84,29% maka dapat disimpulkan bahwa kualitas prototipe yang dibangun adalah Sangat Baik.

## 5. KESIMPULAN

Dari proses penelitian, dapat disimpulkan bahwa penambahan atribut *Quantity* pada analisis data pelanggan terbukti membentuk model klasterisasi yang lebih optimum, menggunakan analisis RFM termodifikasi yaitu QRFM. Klaster optimum diperoleh melalui enam kali percobaan pada dua segmen pelanggan berdasarkan nilai *Davies Bouldin Index (DBI)* sebesar 0,527 pada atribut *QRF (Quantity, Recency dan Frequency)* yang dianalisis menggunakan algoritme K-Means. Dengan demikian, faktor *Quantity* dapat menjadi faktor pembentuk klaster yang lebih baik daripada model orisinal RFM.

Pada masa mendatang, studi ini dapat dikembangkan dengan tantangan mengetahui segmentasi pelanggan berdasarkan metode pembelian yang terdiri dari jenis transaksi *online* maupun *offline*. Analisis juga dapat dilakukan pada jenis produk yang sama dan dibeli berulang pada pembelian berikutnya. Hal-hal tersebut menarik untuk dianalisis sebagai strategi pemasaran baru, mengingat penetrasi penjualan sektor ritel secara online di Indonesia meningkat pesat dalam satu dekade terakhir, dan diduga akan mengubah pola serta metode pembelian.

## DAFTAR PUSTAKA

- BUTARBUTAR, N., PERDANA WINDARTO, A., HARTAMA, D. & SOLIKHUN, 2016. Komparasi Kinerja Algoritma Fuzzy C-Means dan K-Means dalam Pengelompokan Data Siswa Berdasarkan Prestasi Nilai Akademik Siswa (Studi Kasus: SMP Negeri 2 Pematangsiantar). *Jurnal Riset Informasi & Teknis Informatika*, 1(1).
- DYSYANDI, W., SUMARYONO, W., WIDYASTUTI, S. & LESMANA, H., 2019. Bauran Pemasaran Tentang Konsep Apotek Modern Serta Strategi Pemasarannya. *Jurnal Riset Bisnis*, 3(1), hal.1–8.
- FEBRIAN, R., DZULFAQOR, F., LESTARI, M.N., ROMADHON, A.A. & WIDODO, E., 2018. Analisis pola pembelian obat di apotek uii farma menggunakan metode algoritma apriori. In: *Prosiding Seminar Nasional Teknologi Informasi dan Multimedia*. hal.49–54.
- HENDRAWAN, R.A., UTAMIMA, A. & HUSNA, A., 2015. Segmentasi dan Evaluasi Loyalitas Pelanggan Distributor Produk Etikal Farmasi Berdasarkan Nilai Pelanggan. In: *Seminar Nasional Sistem Informasi Indonesia*. hal.2–3.
- HIDAYATULLAH, D.P., ROKHMAWATI, R.I. & PERDANAKUSUMA, A.R., 2018. Analisis Pemetaan Pelanggan Potensial Menggunakan Algoritma K-Means dan LRFM Model Untuk Mendukung Strategi Pengelolaan Pelanggan ( Studi Pada Maninjau Center Kota Malang ). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(8), hal.2406–2415.
- HUGHES, A.M., 1994. *Strategic database marketing*. Chicago: Probus Publishing.
- JAJULI, M. & SINURAYA, R.K., 2018. Artikel Tinjauan: Faktor-faktor yang Mempengaruhi dan Risiko Pengobatan Swamedikasi. *Jurnal Farmaka*, 16(1), hal.48–53.
- LUO, Y., CAI, Q.R., XI, H.X., LIU, Y.J. & YU, Z.M., 2012. Telecom customer segmentation with K-Means clustering. *ICCSE 2012 - Proceedings of 2012 7th International Conference on Computer Science and Education*, (Iccse), hal.648–651.
- RAMADHAN, A., EFENDI, Z. & MUSTAKIM, 2017. Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling. *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 9, Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau*, hal.219–226.
- RESTIYONO, A., 2016. Analisis Faktor yang Berpengaruh dalam Swamedikasi Antibiotik pada Ibu Rumah Tangga di Kelurahan Kajen Kabupaten Pekalongan. *Jurnal Promosi Kesehatan Indonesia*, 11(1), hal.14.
- SUYANTO, 2017. *Data Mining untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika.
- SYAFITRI, I.N., HIDAYATI, I.R. & PRISTIANTY, L., 2017. *Jurnal Farmasi dan Ilmu Kefarmasian Indonesia* Vol. 6 No. 1 Juli 2019 1. *Jurnal Farmasi Dan Ilmu Kefarmasian Indonesia*, 4(1), hal.19–24.
- TRIYANTO, W.A., 2015. Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran. *Simetris*, 6(1), hal.183–188.
- WIDIARINA, W. & ROMI SATRIA WAHONO, 2015. Algoritma Cluster Dinamik Untuk Optimasi Cluster Pada Algoritma K-Means Dalam Pemetaan Nasabah Potensial Algoritma Cluster Dinamik Untuk Optimasi Cluster Pada Algoritma K-Means Dalam. *Journal of Intelligent Systems*, 1(1), hal.33–36.
- YOHANES, 2016. Analisis Perbandingan Algoritma Fuzzy C-Means dan K-Means. 2(1), hal.151–155.

*Halaman ini sengaja dikosongkan*