

## ANALISIS LOYALITAS PELANGGAN BERBASIS MODEL *RECENCY*, *FREQUENCY*, DAN *MONETARY* (RFM) DAN *DECISION TREE* PADA PT. SOLO

Basri\*<sup>1</sup>, Windu Gata<sup>2</sup>, Risnandar<sup>3</sup>

<sup>1,2</sup> Fakultas Ilmu Komputer, STMIK Nusa Mandiri, <sup>2</sup> Pusat Penelitian Informatika-LIPI  
Email: <sup>1</sup>fiogemilang14@gmail.com, <sup>2</sup>windu.gata@gmail.com, <sup>3</sup>risnandar01@gmail.com

\*Penulis Korespondensi

(Naskah masuk: 30 Juli 2019, diterima untuk diterbitkan: 05 Oktober 2020)

### Abstrak

Perkembangan bisnis alat tulis kantor dan sekolah saat ini banyak yang menjanjikan, maka banyak bermunculan pemasok baru dalam bisnis Alat Tulis Kantor dan Sekolah (ATKS). PT Solo yang bergerak di bidang bisnis ATKS harus memiliki strategi dalam setiap persaingan usaha, khususnya dalam meraih loyalitas pelanggan. Loyalitas pelanggan sering dipengaruhi oleh faktor jumlah aktivitas transaksi, nilai nominal transaksi, waktu transaksi di perusahaan, dan atribut *outlet*. Penelitian ini mengusulkan model *Recency*, *Frequency*, dan *Monetary* (RFM) yang dikombinasikan dengan *Decision Tree*. Model RFM digunakan untuk proses klusterisasi data pelanggan berdasarkan jumlah transaksi, nilai nominal transaksi, waktu transaksi, dan atribut outlet. Sedangkan *Decision Tree* dapat menggambarkan tingkat loyalitas pelanggan. Data transaksi dalam penelitian ini dilakukan sepanjang 1 Januari hingga 31 Desember 2018 terhadap 1.203 pelanggan dan 18.087 transaksi melalui faktur pembelian. Hasil penelitian ini menunjukkan bahwa *state-of-the-art* pada model RFM dan *Decision Tree* yang diusulkan lebih unggul dibandingkan hanya dengan menggunakan model RFM saja. *Cluster* ke-1 memiliki 860 pelanggan menghasilkan loyalitas pelanggan sedang (biru), *cluster* ke-2 memiliki 69 pelanggan menghasilkan loyalitas pelanggan yang tinggi (hijau), dan *cluster* ke-3 memiliki 274 pelanggan menghasilkan loyalitas pelanggan yang rendah (merah). Model klusterisasi RFM dan klasifikasi *Decision Tree* telah menghasilkan atribut *outlet* yang berpengaruh terhadap nilai akurasi sebesar 67,54%.

**Kata kunci:** *RFM*, *decision tree*, klasifikasi, loyalitas pelanggan

## CUSTOMER LOYALTY ANALYSIS BASED ON THE *RECENCY*, *FREQUENCY* AND *MONETARY* (RFM) AND *DECISION TREE* MODELS AT PT. SOLO

### Abstract

The development of office and school stationery business at this time, many promising, so many new suppliers have sprung up in the office and school stationery business. PT Solo, which has the office and school stationery business, must have a strategy in every business competition, especially in achieving customer loyalty. Customer loyalty is often influenced by factors in the number of transaction activities, transaction nominal value, transaction time at the company, and outlet attributes. This research proposes a *Recency*, *Frequency*, and *Monetary* (RFM) model combined with a *Decision Tree*. RFM model is used to process customer data clustering based on number of transactions, transaction nominal value, transaction time, and outlet attributes. Whereas *Decision Tree* can describe the level of customer loyalty. Transaction data in this study were conducted from 1 January to 31 December 2018 to the 1,203 customers and 18,087 transactions through purchase invoices. The results of this study indicate that the *state-of-the-art* in the proposed RFM and *Decision Tree* models is outperform compared to only using the RFM model. Cluster 1 has 860 customers resulting in moderate customer loyalty (blue), Cluster 2 has 69 customers resulting in high customer loyalty (green), and Cluster 3 has 274 customers resulting in lower customer loyalty (red). RFM clustering model and *Decision Tree* classification have produced outlet attributes that affect the accuracy value of 67.54%.

**Keywords:** *RFM*, *decision tree*, classification, customer loyalty

### 1. PENDAHULUAN

Perkembangan dunia usaha saat ini semakin bersaing dengan banyaknya perusahaan perusahaan baru yang muncul, tidak terkecuali perusahaan di bidang Alat Tulis Kantor dan Sekolah (ATKS).

Bisnis usaha di bidang ATKS meliputi bidang usaha pengadaan atau pemasok produk alat tulis kantor dan sekolah yang akan memasok produk ke penjual ATKS. Peluang usaha di bidang pengadaan atau pemasok alat tulis kantor dan sekolah terlihat

sangat menjanjikan, maka banyak pengusaha yang membuka usaha di bidang ini dan menimbulkan persaingan. Persaingan yang ketat antar perusahaan pemasok ATKS harus diperhatikan dengan serius oleh pemilik perusahaan pemasok, sehingga perusahaan yang dijalankan tidak tersaingi oleh perusahaan lain yang bergerak di bidang yang sama.

Salah satu strategi bisnis yang dapat dipergunakan oleh perusahaan pemasok ATKS adalah pemanfaatan teknologi informasi sebagai media untuk mengumpulkan data transaksi. Pelanggan merupakan aset perusahaan yang perlu dijaga didalam suatu bisnis, sehingga perusahaan harus memiliki cara supaya aset tersebut bisa lebih berkembang (Li et al., 2018). Sebagaimana yang sudah dijelaskan (Kandeil, Saad, dan Youssef, 2014) bahwa teknologi informasi dalam dunia *e-commerce* dapat dijadikan salah satu sistem rekomendasi berdasarkan penilaian pelanggan. Dengan menerapkan manajemen pelanggan maka perusahaan dapat mengetahui data pelanggan yang memiliki loyalitas tinggi, sedang, dan kurang/rendah.

Metode *clustering* bertujuan untuk memprediksi kelas target untuk setiap *record* data baru berupa data yang tidak dalam data historis (Sudriyanto, 2017). *Clustering* juga digunakan untuk menganalisis kelompok berdasarkan sektor penjualan, dan lokasi pelanggan, sebagai cara untuk meningkatkan analisis pengelompokan (Kandeil, Saad, and Youssef 2014). Sumber yang dapat dipergunakan untuk mengolah keterkaitan pelanggan dengan perusahaan adalah data pelanggan dan perkembangan teknologi (Ngai, Xiu, dan Chau, 2009). Metode CRM akan menghasilkan klasifikasi dan klastering, klasifikasi akan menghasilkan kelompok pelanggan berdasarkan geografis dan jarak dengan perusahaan (Qi dan Zhang, 2012). Sedangkan *clustering* akan menghasilkan klaster pelanggan berdasarkan hubungan antara pelanggan dengan perusahaan (Cho et al. 2012), (Sudriyanto, 2017), (Kandeil, Saad, dan Youssef, 2014).

Data pelanggan cenderung merupakan data transaksi pelanggan yang memiliki banyak atribut, sehingga mengakibatkan pengolahan data yang dilakukan akan menjadi kurang baik (Hu dan Yeh, 2014). Data pelanggan yang memiliki banyak atribut perlu dilakukan pemilihan atribut yang baik atau yang cocok untuk menghasilkan *output* yang baik. Pemilihan atribut yang akan digunakan dalam penelitian ini mengacu pada metode *Recency*, *Frequency*, dan *Monetary* (RFM) (Kandeil, Saad, and Youssef 2014), (Ait Daoud et al. 2016), (Maryani dan Riana, 2017), dan (Sheshasaayee dan Logeshwari, 2019).

*Recency* merupakan sebuah nilai yang didapat dari jarak seorang pelanggan terakhir melakukan pembelian produk pada perusahaan sampai data yang diteliti, satuan yang sering dipergunakan untuk

menghitung *Recency* adalah hari, bulan, dan tahun, semakin dekat dengan dengan pengambilan data maka pelanggan tersebut memiliki *point* lebih tinggi (Sudriyanto, 2017). *Frequency* adalah seberapa sering pelanggan melakukan transaksi kepada organisasi (Sudriyanto, 2017).

Algoritma K-Means merupakan teknik dalam data mining untuk mengelompokkan objek data berdasarkan kedekatan karakteristik (Reza Allahyari Soeini dan Ebrahim Fathalizade, 2012). Algoritma K-Means menggunakan *input* yang menyatakan jumlah *cluster* sebanyak *k* dan merupakan proses penentuan kelompok objek berdasarkan jarak terdekat yang dilakukan secara iteratif. Bagaian penting dari algoritma K-Means adalah titik tengah sebagai pusat *cluster* yang disebut sebagai *centroid* (Chrisnanto dan Kanianingsih, 2019). Nilai *centroid* awal diambil secara acak sesuai dengan jumlah *k* yang sudah ditentukan dan selanjutnya menghitung jarak setiap objek terhadap semua *centroid* dan mengelompokkan objek dengan *centroid* yang terdekat dengan objek, proses menghitung dan mengelompokkan ini akan dilakukan sampai semua objek tidak berubah posisi terhadap *centroid*. Jika objek sudah tidak berubah posisi, maka dapat ditentukan *cluster* dari data yang dianalisis tersebut. Algoritma K-Means juga merupakan algoritma klasik untuk memecahkan masalah *clustering*. Metode K-Means juga sering disebut dengan metode non-hirarkis. Metode ini merupakan teknik pendididkan (*partition*) yang membagi atau memisahkan objek *k* ke daerah terpisah.

Pada analisis sebelumnya yang dilakukan oleh Yulison Herry dan Ade Kanianingsih yang berjudul “Pengelompokan Ekuitas Pelanggan berbasis *recency frequency Monetary* (RFM) menggunakan K-Means *Clustering*”. Mereka melakukan analisis data pelanggan di PT. Berkah Lestari yang menganalisis data selama satu tahun transaksi pada tahun 2018 dengan jumlah transaksi sebanyak 20.814 *invoice* yang dilakukan oleh 894 pelanggan dengan faktur pembelian sebanyak 1.117 *invoice*. Kemudian dilakukan *clustering* sebanyak empat *cluster* dengan karakteristik 23,37% (209 pelanggan) berada di *cluster* 1 dengan indikasi loyalitas sedang, 15,43% (138 pelanggan) berada di *cluster* 2 dengan indikasi tingkat rendah, 20,04% (215 pelanggan) berada di *cluster* 3 dengan indikasi loyalitas baik, dan 22,70% (203 pelanggan) berada di *cluster* 4 dengan indikasi loyalitas pelanggan rendah. Maka dengan metode RFM dengan algoritma K-Means dapat melihat segmentasi loyalitas pelanggan terhadap perusahaan (Chrisnanto dan Kanianingsih, 2019).

Penelitian yang dilakukan oleh Beta Estri Adiana, Indah Soesanti, dan Andhistya Erna Permanasari yang berjudul “Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik *Clustering*”. Penelitian ini dilakukan di UD. Gemilang Kencana, sebuah perusahaan yang

bergerak di bidang usaha makanan pengolahan buah carica menjadi manisan. Analisis dilakukan mulai dari 1 Juni hingga 31 Desember 2017 yang diperoleh data sebanyak 127 data pelanggan. Tujuan dari penelitian ini adalah melakukan segmentasi dan mengukur kesetiaan pelanggan terhadap sebuah produk UKM dengan mengkombinasikan model RFM dan teknik *clustering* menggunakan metode K-Means. Hasil yang diperoleh dari penelitian ini terdapat 3 *cluster* yaitu 30 pelanggan di kelompokkan sebagai kategori *typical customers*, 8 pelanggan dikategorikan sebagai kategori *superstar* dan 89 pelanggan dikategorikan sebagai pelanggan *dormant customer*. Pelanggan dapat mengidentifikasi pelanggan dengan nilai tinggi dan berpotensi memberi keuntungan serta dapat membantu dalam merancang strategi pemasaran yang tepat untuk tiga *cluster* yang terbentuk (Li et al. 2018).

Menurut penelitian yang dilakukan oleh Tommi Alfian Armawan Sandi, Mugi Raharjo, Jordy Lasmana Putra, dan Ridwan dengan judul “*Clustering Kesetiaan Pelanggan e-retail dengan Model RFM (Recency, Frequency, Monetary)*” dan K-Means”. Penelitian ini menggunakan *dataset* yang bersumber dari UCI *Repository Retail Online* pada periode 1 Desember 2010 hingga 18 Januari 2011. Penelitian ini menggunakan teknik RFM dan Algoritma K-Means, yang menghasilkan 3 klastering, di mana terdapat 16 pelanggan di klaster 1, 2 pelanggan terdapat di klaster 2, dan 444 pelanggan terdapat di klaster 3 (Sandi et al. 2018).

Sementara itu, Aulia Dewi Savitri, Fitra Abdurrachman Bachtiar, dan Nanang Yudi Setiawan melakukan penelitian serupa yang berjudul “*Segmentasi Pelanggan Menggunakan Metode K-Means Clustering Berdasarkan Model RFM pada Klinik Kecantikan Belle Crown, Malang*”. Data yang digunakan dalam penelitian ini berdasarkan transaksi pelanggan periode Mei-Oktober 2017 dari sejumlah 21.513 transaksi dan 4.716 pelanggan. Penelitian ini menggunakan model RFM dan metode K-Means dengan menguji 2 dan 3 *cluster*. Analisis yang dilakukan dengan 2 *cluster* diperoleh *cluster* 1 adalah *cluster* yang mendudukkan peringkat ke-2. Sedangkan *cluster* 2 menempati peringkat ke-1. Berdasarkan analisis dengan menggunakan 3 *cluster* menghasilkan *cluster* 1 menempati peringkat ke-1, *cluster* 2 menempati peringkat ke-2, dan *cluster* 3 menempati peringkat ke-3. Penelitian ini telah mendeskripsikan profil atau karakteristik setiap segmen, dibuktikan oleh nilai *recency* untuk mengetahui rentang waktu transaksi terakhir yang dilakukan oleh pelanggan, nilai *frequency* untuk mengetahui jumlah transaksi pelanggan, dan nilai *monetary* untuk mengetahui besaran Rupiah yang dikeluarkan pelanggan untuk mendapatkan layanan di Belle Crown Malang (Savitri, Bachtiar, dan Setiawan, 2018).

Menurut penelitian yang terdahulu yang dilakukan oleh Daniel Pradipta Hidayatullah, Retno

Indah Rochmawati, Andi Reza Perdanakusuma yang berjudul “*Analisis Pemetaan Pelanggan Potensial Menggunakan Algoritma K-Means dan LRFM Model untuk Mendukung Strategi Pengelolaan Pelanggan*”. Penelitian ini dilakukan di Maninjau Center, Kota Malang. Data yang dianalisis berupa data pelanggan tahun 2016 menggunakan parameter LRFM (*Length, Recency, Frequency, Monetary*) dan menggunakan algoritma K-Means dengan *clustering* yang diinginkan adalah 4. Hasil yang didapat dari studi kasus ini telah menunjukkan bahwa segmen pelanggan yang terbentuk ada 4 *cluster*, yaitu *cluster* 1 berjumlah 2, *cluster* 2 berjumlah 41, *cluster* 3 berjumlah 3, dan *cluster* 4 berjumlah 657. Hasil dari penelitian ini akan menghasilkan strategi marketing yang harus sesuai dengan karakteristik pelanggan di setiap *clustering* yang terbentuk (Hidayatullah<sup>1</sup>, Rokhmawati<sup>2</sup>, dan Perdanakusuma<sup>3</sup> 2018).

Berdasarkan permasalahan yang sudah dibahas berdasarkan penelitian-penelitian sebelumnya, maka penelitian ini akan memberikan kontribusi baru berupa model *Recency, Frequency, dan Monetary* (RFM) yang dikombinasikan dengan *Decision Tree*. Model RFM digunakan untuk proses klasterisasi data pelanggan berdasarkan jumlah transaksi, nilai nominal transaksi, waktu transaksi, dan atribut outlet. Sedangkan *Decision Tree* dapat menggambarkan tingkat loyalitas pelanggan. Data transaksi dalam penelitian ini dilakukan sepanjang 1 Januari hingga 31 Desember 2018 terhadap 1.203 pelanggan dan 18.087 transaksi melalui faktur pembelian. Hasil penelitian ini menunjukkan bahwa *state-of-the-art* pada model RFM dan *Decision Tree* yang diusulkan lebih unggul dibandingkan hanya dengan menggunakan model RFM saja sebagaimana penelitian sebelumnya yang paling mendekati dengan topik ini (Mahboubeh et al. 2011) dan (Li et al. 2018).

## 2. METODE PENELITIAN

### 2.1. Pengumpulan Data

Data yang akan dianalisis dalam penelitian ini bersumber dari data perusahaan PT. Solo, Jakarta yang bertempat di Jakarta. Pada tahap ini, data dikumpulkan berupa data pelanggan yang bertransaksi di perusahaan PT. Solo, Jakarta, di antaranya data waktu transaksi yang dilakukan pelanggan, jumlah transaksi yang dilakukan oleh pelanggan, dan data kapan pelanggan melakukan transaksi terakhir.

### 2.2. Pra-pemrosesan Data

Data PT. Solo Jakarta merupakan data mentah, sehingga dibutuhkan pra-pemrosesan data kembali. Bila menghasilkan nilai kosong atau nilainya tidak benar, maka akan dibersihkan pada pra-pemrosesan data ini. Selanjutnya, memilih atribut yang dibutuhkan dari 12 atribut yang terdapat pada data

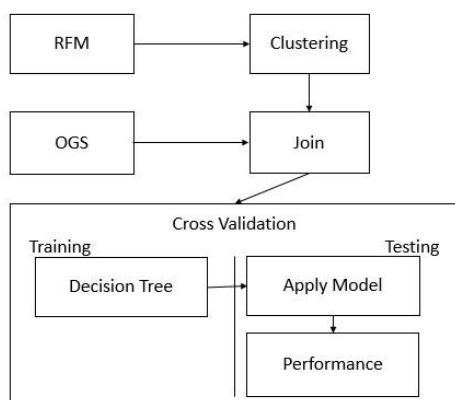
awal, dengan cara menghilangkan atribut yang yang dibutuhkan.

### 2.3. Model RFM

Model *Recency*, *Frequency*, dan *Monetary* (RFM) melakukan ekstraksi data dengan atribut yang dibutuhkan untuk mengukur nilai loyalitas pelanggan. Data di sini sudah melewati tahap pengumpulan data dan pra-pemrosesan data. Model RFM akan melakukan proses klasterisasi yang terdiri dari *Recency* untuk menghitung tanggal terakhir transaksi yang dilakukan oleh pelanggan pada periode 2018, *Frequency* untuk menghitung jumlah transaksi pelanggan selama periode 2018, dan *Monetary* untuk menghitung nilai atau total transaksi pelanggan selama periode 2018 dalam bentuk uang. Skor RFM dapat dihitung sebagai berikut:

Skor RFM = ((nilai *Recency* x 100) + (nilai *Frequency* x 0) + nilai *Monetary*.....(1)

### 2.4. Algoritma



Gambar 1 . Model Klasifikasi Berdasarkan Klasterisasi

Data yang sudah dihasilkan oleh proses model RFM dilakukan proses *clustering*. Selanjutnya dilakukan proses *join* dengan data *Outlet*, Gudang, dan Sales (OGS). Penggabungan data RFM dengan OGS, diperlukan proses *Cross Validation* dengan memasukkan *data training* dan *data testing*, sehingga menghasilkan data yang sudah diklasifikasi.

## 3. HASIL

### 3.1. Pengumpulan Data

Data uji yang dikumpulkan pada penelitian ini bersumber dari *database* penjualan PT. Solo, Jakarta, yang bergerak di bidang pemasok produk ATKs di wilayah Jakarta. Data pelanggan di PT. Solo memiliki atribut : *Customer Id*, *Customer Name*, *Kota*, *Invoice no*, *Month*, *Detail Date*, *Salesman*, *Item Nomer*, *Item Description*, *Quantity*, *Carton*, *Sales Amount*, *Item Prince*, *Discount 1*, *Discount 2*, *Discount 3*, *Discount 4*, *Item Category*, *type*, *ID*, Gudang, *Outlet*, dan Nomor PO.

Data yang akan dianalisis dari PT. Solo, Jakarta ini di ambil selama transaksi satu tahun yang dimulai

dari tanggal 1 Januari 2018 hingga 31 Desember 2018, sejumlah 24 atribut dan 81.928 *record*.

Tabel 1. Data mentah transaksi pelanggan

No	Nama Atribut	Keterangan
1	Customer Id	Kode unit pelanggan yang sudah pernah melakukan transaksi
2	Customer Name	Nama Pelanggan yang sudah pernah bertransaksi
3	Kota	Kota tempat alamat pelanggan
4	Invoice	Nomor unik transaksi penjualan
5	Month	Bulan transaksi penjualan
6	Detail Date	Tanggal transaksi penjualan
7	Selesman	Sales dari perusahaan
8	Item No	Kode unit produk yang terjual
9	Item Desc	Nama barang yang terjual
10	Quantity	Jumlah barang yang terjual pada setiap transaksi
11	Sales Amount	Harga barang perunit di kali jumlah barang yang terjual
12	Item Price	Harga masing masing barang

### 3.2. Pra-pemrosesan Data

Tahapan dari pra-pemrosesan data dimulai dari proses pengumpulan data pelanggan dari PT. Solo, Jakarta pada periode 1 Januari 2018 hingga 31 Desember 2018 yang masih berbentuk data mentah, di mana pra-pemrosesan data akan menghilangkan atribut-atribut yang tidak dibutuhkan.

Tabel 2. Data proposes

No	Nama Atribut	Keterangan
1	Customer Id	Kode unit pelanggan yang sudah pernah melakukan transaksi
2	Customer Name	Nama Pelanggan yang sudah pernah bertransaksi
3	Invoice	Nomor unik transaksi penjualan
4	Detail Date	Tanggal transaksi penjualan
5	Sales Amount	Harga barang perunit di kali jumlah barang yang terjual

### 3.3. Atribut Model RFM

Memilih atribut yang dibutuhkan untuk penelitian ini dengan menggunakan metode RFM dalam periode pelanggan melakukan transaksi terakhir dengan periode analisis, jumlah frekuensi transaksi yang dilakukan pelanggan selama periode yang di analisis, dan jumlah nominal transaksi untuk setiap pelanggan dari PT. Solo, Jakarta selama periode yang dianalisis.

Nilai kriteria R (*recency*) yang membutuhkan atribut untuk menunjukkan adanya rentang waktu transaksi terakhir pelanggan dengan periode analisis, sehingga atribut yang dibutuhkan adalah atribut *Detail Date* dan *Customer Id*, tanggal transaksi akhir yang dilakukan oleh pelanggan di perusahaan PT. Solo Jakarta.

Nilai kriteria F (*frequency*) yang memerlukan atribut untuk menghitung jumlah transaksi pelanggan di PT. Solo Jakarta. Atribut yang dipergunakan betupa *invoice* dan *Customer Id*. Setiap

pelanggan akan dihitung berapa jumlah transaksi pada rentang waktu penelitian.

Nilai kriteria M (*Monetary*) yang membutuhkan atribut untuk menunjukkan total nominal transaksi setiap pelanggan, atribut yang diperlukan berupa *Sales Amount* dan *Customer Id*

### 3.4. Hasil Model RFM

Setelah melakukan pemilihan atribut yang terdapat pada data awal penelitian maka didapat nilai *Recency*, *Frequency*, dan *Monetary* seperti yang terlihat pada tabel 3.

Tabel 3. Data cuplikan transformasi data ke dalam RFM

No	Customer Id	R	F	M
1	TJBGG7UTM	144	3	4254896
2	TJBGGADJY	17	67	136694792
3	TJBGGAFTA	10	15	49781151
4	TJBGGALAM	13	34	35121923
5	TJBGGALIF	67	24	82991221
6	TJBGGALII	74	42	107260820
7	TJBGGALM6	14	83	343776294
8	TJBGGANNR	12	17	34776123
9	TJBGGARIS	32	18	39601794
10	TJBGGARVN	101	21	185208442
11	TJBGGATAK	49	8	4805643
12	TJBGGBARU	19	48	98326440
13	TJBGGBGST	10	37	526053452
14	TJBGGBKJY	21	28	66656076
15	TJBGGBRIT	286	6	13738661
16	TJBGGBRMS	34	43	139889453
17	TJBGGBTPL	248	4	10150152
18	TJBGGBTM	175	3	1280816
....				
1203	TJTPGLOTS	73	15	170713033

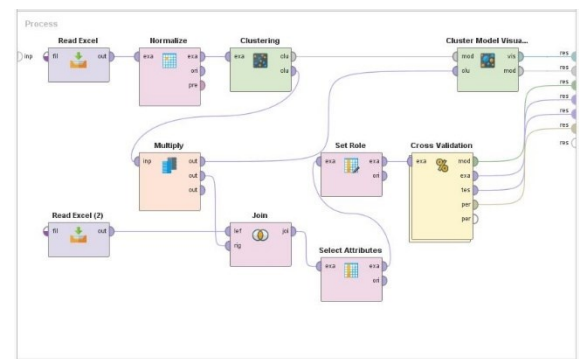
Setelah diperoleh nilai RFM, jika masih terlihat perbedaan nilai yang sangat tinggi, maka itu perlu dilakukan normalisasi. Dengan demikian, nilai RFM memiliki perbedaan nilai yang tidak tinggi seperti yang terlihat pada tabel 4.

Tabel 4. Data Cuplikan Transformasi Data pada RFM Setelah Proses Normalisasi

No	Customer Id	R	F	M
1	TJBGG7UTM	0.801	-0.673	-0.421
2	TJBGGADJY	-0.782	2.935	0.675
3	TJBGGAFTA	-0.870	0.002	-0.044
4	TJBGGALAM	-0.832	10.743	-0.166
5	TJBGGALIF	-0.158	0.510	0.230
6	TJBGGALII	-0.071	1.525	0.431

No	Customer Id	R	F	M
7	TJBGGALM6	-0.820	3.837	23.900
8	TJBGGANNR	-0.845	0.115	-0.168
9	TJBGGARIS	-0.595	0.172	-0.129
10	TJBGGARVN	0.265	0.341	10768
11	TJBGGATAK	-0.383	-0.391	-0.417
12	TJBGGBARU	-0.757	18.637	0.357
13	TJBGGBGST	-0.870	12.434	3.899
14	TJBGGBKJY	-0.732	0.735	0.095
15	TJBGGBRIT	25.739	-0.504	-0.343
16	TJBGGBRMS	-0.570	15.818	0.701
17	TJBGGBTPL	2.099	-0.617	-0.372
18	TJBGGBTM	11.888	-0.673	-0.446
....				
1203	TJTPGLOTS	-0.084	0.003	0.957

### 3.5. Proses algoritma



Gambar 2. Design proses algoritma klasifikasi berdasarkan klusterisasi

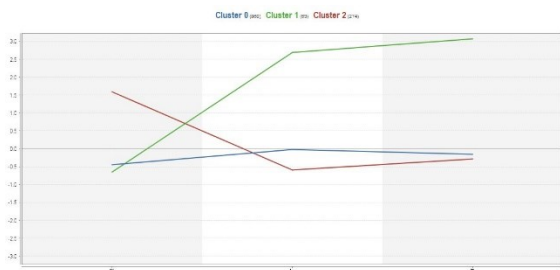
Pada proses klusterisasi ini, digunakan algoritma K-Mean terlebih dahulu untuk mengetahui jumlah k yang menyatakan banyaknya segmen data yang diinginkan. Dalam hal ini k yang dipilih adalah 3, dengan pertimbangan segmentasi yang dilakukan sebanyak 3 kelompok karakteristik pelanggan dalam melakukan transaksi.

Proses *clustering* dengan algoritma K-Mean akan menggunakan RapidMiner yang diawali dengan membaca file data yang sudah dilakukan pemilihan atribut RFM yang menggunakan format excel. Selanjutnya proses normalisasi untuk mendapatkan nilai *Recency*, *Frequency*, dan *Monetary* yang sudah diukur supaya nilai tersebut berada dalam *range* yang kecil. Setelah proses normalisasi, selanjutnya dilanjutkan proses *clustering* data, sehingga pada proses ini terbentuk kelompok pelanggan menjadi 3 kelompok.

Tabel 5. *Centroid table* merupakan hasil pengolahan yang dilakukan dari perhitungan titik pusat data atas setiap anggota masing-masing *cluster* yang sudah terbentuk

Tabel 5. Centroid Table

Attribute	Cluster_0	Cluster_1	Cluster_2
Recency (R)	-0.453	-0.653	1.585
Frequency (F)	-0.027	2.696	-0.593
Monetary (M)	-0.153	3.071	-0.292



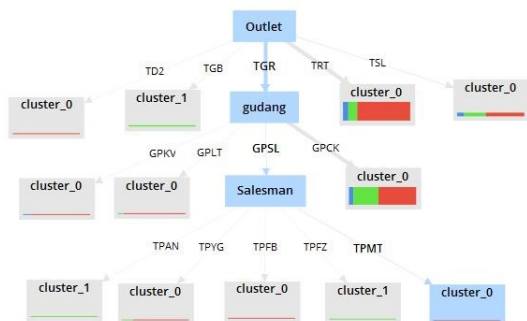
Gambar 3. Visualization klusterisasi

Gambar 3 menjelaskan tampilan grafik yang terbagi menjadi 3 kelompok atau *cluster* yang berwarna biru merupakan *cluster* ke-1 yang memiliki loyalitas sedang, yang berwarna hijau merupakan *cluster* ke-2 yang memiliki loyalitas tinggi dan yang berwarna merah adalah *cluster* ke-3 yang memiliki loyalitas rendah.

Tabel 6. Cluster Model

Cluster	Jumlah
1	860
2	69
3	274
Total	1203

Tabel 6 menunjukkan jumlah *cluster* dan jumlah anggota yang terdapat pada masing-masing *cluster* yaitu *cluster* ke-1 terdapat 860 pelanggan, *cluster* ke-2 terdapat 69 pelanggan, dan *cluster* ke-3 terdapat 274, sehingga jika ditotalkan semua akan berjumlah 1.203 atau sejumlah pelanggan yang terdapat pada data yang dianalisis. Sementara dari gambar 4 terlihat bahwa atribut sangat berpengaruh besar pada data adalah atribut outlet, setelah itu atribut gudang dan urutan berikutnya adalah atribut salesman.



Gambar 4. Visualisasi Decision Tree

#### 4. KESIMPULAN

Penelitian ini telah menunjukkan bahwa kelompok pelanggan PT. Solo, Jakarta dalam bertransaksi selama satu tahun pada 2018, dari jumlah

transaksi sebanyak 79.225 transaksi yang dilakukan oleh 1.203 pelanggan dengan faktor pembelian sebanyak 18.087. Hasil analisis yang dilakukan adalah pelanggan yang terdapat pada kelompok *cluster* ke-1 merupakan pelanggan memiliki loyalitas yang sedang, pelanggan yang terdapat pada *cluster* ke-2 merupakan kelompok pelanggan yang memiliki loyalitas yang tinggi terhadap perusahaan, dan *cluster* ke-3 merupakan kelompok pelanggan yang memiliki loyalitas yang rendah.

Hasil penelitian ini dapat digunakan oleh PT. Solo, Jakarta untuk membuat langkah strategis kepada *cluster* ke-1 untuk menaikkan jumlah nominal transaksi, menghubungi pelanggan di *cluster* ke-2 untuk kembali berbelanja sehingga *recency* naik dan membuat strategi agar pelanggan yang berada di *cluster* ke-3 dapat memperbanyak jumlah transaksi.

Perlunya melakukan evaluasi terhadap outlet TD2, TRT, dan TSL karena terdapat *cluster* ke-1 dengan loyalitas sedang, dan dinilai perlu untuk mengevaluasi kinerja salesman TPYG, TPFB, TPMT.

Dari penelitian yang dilakukan dengan metode klasifikasi berdasarkan klasifikasi pelanggan dengan model RFM yang dikombinasikan dengan *Decision Tree* model, dapat menghasilkan nilai akurasi sebesar 67,54 %.

#### DAFTAR PUSTAKA

- AIT DAOUD, dkk. 2016. Combining RFM Model and Clustering Techniques for Customer Value Analysis of a Company Selling Online. In *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*.
- CHO, YOUNG SUNG, dkk. 2012. Implementation of Personalized Recommendation System Using K-Means Clustering of Item Category Based on RFM. In *2012 IEEE 6th International Conference on Management of Innovation and Technology, ICMIT 2012*.
- CHRISNANTO, dkk. 2019. Pengelompokan Ekuitas Pelanggan Berbasis Recency Frequency Monetary ( RFM ) Menggunakan K-Means Clustering. 2019(Sentika): 13–14.
- HIDAYATULLAH, dkk. 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 (J-PTIIK) Universitas Brawijaya* 2(8): 2548–2964. <http://j-ptiik.ub.ac.id>.
- HU, YA HAN, & TZU WEI YEH. 2014. Discovering Valuable Frequent Patterns Based on RFM Analysis without Customer Identification Information. *Knowledge-Based Systems* 61: 76–88.

- <http://dx.doi.org/10.1016/j.knosys.2014.02.009>.
- KANDEIL, DALIA ABDELRAZEK, dkk. 2014. A Two-Phase Clustering Analysis for B2B Customer Segmentation. In *Proceedings - 2014 International Conference on Intelligent Networking and Collaborative Systems, IEEE INCoS 2014*,.
- LI, ZHIBIN, dkk. 2018. Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering. *Jutei* 2(2): 23–32.
- MAHBOUBEH, KHAJVAND dkk. 2011. Estimating Customer Lifetime Value Based on RFM Analysis of Customer Purchase Behavior: Case Study. In *Procedia Computer Science 3 (2011): 57-63*.
- MARYANI, INA, & DWIZA RIANA. 2017. Clustering and Profiling of Customers Using RFM for Customer Relationship Management Recommendations. In *2017 5th International Conference on Cyber and IT Service Management, CITSM 2017*,.
- NGAI, E. W.T., LI XIU, & D. C.K. CHAU. 2009. Application of Data Mining Techniques in Customer Relationship Management: A Literature Review and Classification. *Expert Systems with Applications* 36(2 PART 2): 2592–2602.  
<http://dx.doi.org/10.1016/j.eswa.2008.02.021>.
- QI, LILEI, & SONGJUAN ZHANG. 2012. The Development of Customer Relationship Management System Based on Rough Set. *Communications in Computer and Information Science* 315: 328–33.
- REZA ALLAHYARI SOEINI, & EBRAHIM FATHALIZADE. 2012. Customer Segmentation Based on Modified RFM Model in Insurance Industry. *Ipcsit* 25: 101–4.  
[www.ipcsit.com/vol25/020-ICMLC2012-L0071.pdf](http://www.ipcsit.com/vol25/020-ICMLC2012-L0071.pdf).
- SANDI, TOMMI ALFIAN ARMAWAN, dkk. 2018. Clustering Kesetiaan Pelanggan E-Ritel Dengan Model Rfm. *PILAR Nusa Mandiri* 14(2): 239–46.
- SAVITRI, AULIA DEWI, dkk. 2018. Segmentasi Pelanggan Menggunakan Metode K-Means Clustering Berdasarkan Model RFM Pada Klinik Kecantikan (Studi Kasus : Belle Crown Malang). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya* 2(9): 2957–66.
- SHESHASAYEE, ANANTHI, & L. LOGESHWARI. 2019. Implementation of RFM Analysis Using Support Vector Machine Model. In *Proceedings of the International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), I-SMAC 2018*.
- SUDRIYANTO. 2017. Prosiding SNATIF Ke-4 Clustering Loyalitas Pelanggan Dengan Metode RFM (Recenty,Frequency,Monetary) Dan Fuzzy C-Means.

*Halaman ini sengaja dikosongkan*