

KOMPARASI METODE *CLUSTERING K-MEANS* DAN *K-MEDOIDS* DENGAN MODEL *FUZZY RFM* UNTUK PENGELOMPOKAN PELANGGAN

Elly Muningsih
Manajemen Informatika, AMIK BSI Yogyakarta
Email : elly.emh@bsi.ac.id

Abstract ~ *The K-Means method is one of the clustering methods that is widely used in data clustering research. While the K-Medoids method is an efficient method used for processing small data. This study aims to compare two clustering methods by grouping customers into 3 clusters according to their characteristics, namely very potential (loyal) customers, potential customers and non potential customers. The method used in this study is the K-Means clustering method and the K-Medoids method. The data used is online sales transaction. The clustering method testing is done by using a Fuzzy RFM (Recency, Frequenty and Monetary) model where the average (mean) of the third value is taken. From the data testing is known that the K-Means method is better than the K-Medoids method with an accuracy value of 90.47%. Whereas from the data processing carried out is known that cluster 1 has 16 members (customers), cluster 2 has 11 members and cluster 3 has 15 members.*

Keywords : *clustering, K-Means method, K-Medoids method, customer, Fuzzy RFM model.*

Abstrak ~ Metode *K-Means* merupakan salah satu metode *clustering* yang banyak digunakan dalam penelitian pengelompokan data. Sedangkan metode *K-Medoids* merupakan metode yang efisien digunakan untuk pengolahan data yang kecil. Penelitian ini bertujuan untuk membandingkan atau mengkomparasi dua metode *clustering* dengan cara mengelompokkan pelanggan menjadi 3 *cluster* sesuai dengan karakteristiknya, yaitu pelanggan sangat potensial (loyal), pelanggan potensial dan pelanggan kurang (tidak) potensial. Metode yang digunakan dalam penelitian ini adalah metode *clustering K-Means* dan metode *K-Medoids*. Data yang digunakan adalah data transaksi penjualan online. Pengujian metode *clustering* yang dilakukan adalah dengan menggunakan model *Fuzzy RFM (Recency, Frequenty dan Monetary)* dimana diambil rata-rata (mean) dari nilai ketiga tersebut. Dari pengujian data diketahui bahwa metode *K-Means* lebih baik dari metode *K-Medoids* dengan nilai akurasi 90,47%. Sedangkan dari pengolahan data yang dilakukan diketahui bahwa *cluster 1* memiliki 16 anggota (pelanggan), *cluster 2* memiliki 11 anggota dan *cluster 3* memiliki 15 anggota.

Kata kunci : *clustering, metode K-Means, metode K-Medoids, pelanggan, model Fuzzy RFM.*

A. PENDAHULUAN

Konsumen atau pelanggan merupakan salah satu faktor penting untuk kemajuan perusahaan atau pelaku usaha. Saat ini pelaku usaha bisa berasal dari berbagai macam sektor, dari yang *offline* maupun *online*. Pelaku usaha *online* yang dikenal dengan nama *E-Commerce* merupakan usaha yang dilakukan menggunakan media elektronik yaitu *internet*. Di tengah persaingan bisnis dewasa ini, *online shop* harus menyadari bahwa hubungan dengan konsumen sangat penting karena menunjang perkembangan dan kelangsungan usahanya. Dengan meningkatkan pemahaman akan kebutuhan konsumen secara individu, pihak manajemen harus bisa mengenali konsumen

terbaiknya sehingga dapat mempertahankan loyalitas konsumennya terhadap perusahaan.

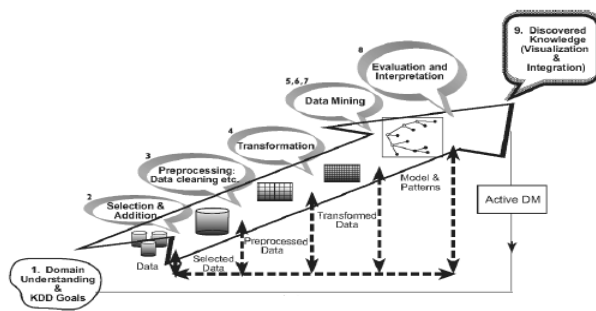
Permasalahan yang terjadi pada pelaku usaha disebabkan karena mereka mengalami kesulitan dalam menentukan pelanggan yang potensial. Menurut Zulkarnain (2012), pelanggan potensial adalah pelanggan yang loyal terhadap perusahaan. Menurut Kotler dalam Zulkarnain (2012), pelanggan potensial memberikan keuntungan yang lebih bagi perusahaan karena salah satu ciri pelanggan potensial adalah merekomendasikan kepada orang lain untuk membeli. Penentuan pelanggan potensial yang dilakukan secara manual akan mendapatkan hasil yang kurang atau tidak akurat.

Metode *K-Means* merupakan metode *clustering* dalam *Data Mining* yang sering digunakan oleh peneliti untuk mengklaster data (Febrianti dkk, 2016). Metode *K-Medoids* merupakan metode *clustering* yang cukup efisien untuk data kecil (Zayuka dkk, 2017). Pada penelitian ini akan dilakukan komparasi atau perbandingan antara metode *K-Means* dan metode *K-Medoids* dengan model *Fuzzy RFM* (*recency, frequency dan monetary*) untuk pengelompokan data pelanggan menjadi 3 klaster. Data yang digunakan adalah data transaksi penjualan online, dan diharapkan hasil penelitian pengelompokan pelanggan ini bisa lebih akurat sehingga dapat memaksimalkan strategi pemasaran yang pada gilirannya dapat meningkatkan penjualan produk.

B. TINJAUAN PUSTAKA

1. Data Mining

Menurut Maimon & Rokach dalam Muningsih (2015), *Data Mining* merupakan sebuah inti dari proses *Knowledge Discovery in Database* (KDD), meliputi dugaan algoritma yang mengeksplor data, membangun model dan menemukan pola yang belum diketahui. KDD bersifat otomatis, dapat didefinisikan sebagai pengorganisasian proses untuk pengidentifikasian yang benar, berguna dan penemuan pola dari kumpulan data yang besar dan kompleks. Tahapan pada proses KDD pada database digambarkan sebagai berikut (Maimon dan Rokach, 2010) :



Gambar II.1 Tahapan proses KDD dalam database
(Sumber : Maimon & Rokach, 2010)

Menurut (Larose, 2005), (Moertini, 2002) dan (Susanto dan Suryadi, 2010) ada enam fungsi utama *data mining*, yaitu :

1. *Description* (deskripsi), untuk memberi gambaran secara ringkas bagi sekumpulan

- data yang jumlahnya sangat besar dan banyak jenisnya. Termasuk dalam fungsi ini adalah metode *Decision Tree, Neural Network, dan Exploratory Data Analysis*.
2. *Estimation* (estimasi), untuk menerka sebuah nilai yang belum diketahui, misal menerka penghasilan seseorang ketika informasi mengenai orang tersebut diketahui. Metode yang digunakan antara lain *Point Estimation dan Confidence Interval Estimations, Simple Linear Regression dan Correlation, dan Multiple Regression*.
3. *Prediction* (prediksi), untuk memperkirakan nilai masa mendatang, misal memprediksi stok barang satu tahun ke depan. Fungsi ini mencakup metode *Neural Network, Decision Tree, dan k-Nearest Neighbor*.
4. *Classification* (klasifikasi), merupakan proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Metode yang digunakan antara lain *Neural Network, Decision Tree, k-Nearest Neighbor, dan Naive Bayes*.
5. *Clustering* (pengelompokan), yaitu pengelompokan mengidentifikasi data yang memiliki karakteristik tertentu. Metode dalam fungsi ini diantaranya *Hierarchical Clustering, metode K-Means, dan Self Organizing Map (SOM)*
6. *Association* (asosiasi), dinamakan juga analisis keranjang pasar dimana fungsi ini mengidentifikasi item-item produk yang kemungkinan dibeli konsumen bersamaan dengan produk lain. Metode atau algoritma dalam fungsi ini adalah *Apriori, Generalized Sequential Pattern (GSP), FP-Growth dan GRI algorithm*

2. Metode K-Means

Menurut Larose (2005) algoritma *K-Means* merupakan sebuah metode sederhana untuk membagi suatu kumpulan data dalam suatu angka spesifik dari *cluster*, yaitu *k*. *K-Means* merupakan suatu metode data *clustering* non hirarki yang mempartisi data ke dalam bentuk satu atau lebih *cluster* atau kelompok, sehingga data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan data yang memiliki karakteristik berbeda dikelompokkan ke dalam kelompok lain (Agusta, 2007). Disebutkan bahwa metode *K-Means* adalah metode yang cepat dan efisien yang

dapat digunakan dalam *clustering* data. Tujuan proses pengelompokan atau *clustering* adalah meminimalkan terjadinya *objective function* yang diset dalam proses *clustering*, yang pada umumnya digunakan untuk meminimalisasikan variasi dalam suatu *cluster* dan memaksimalkan variasi antar *cluster* (Agusta, 2007).

Algoritma dasar *clustering* data menggunakan metode *K-Means* dapat dilakukan dengan cara (Agusta, 2007), (Muningsih, 2015) :

1. Tentukan jumlah *cluster* yang ingin dibentuk
2. Inisialisai *k* sebagai pusat *cluster* (beri nilai-nilai random)
3. Alokasikan data sesuai dengan jumlah *cluster* yang ditentukan. Kedekatan dua obyek ditentukan berdasarkan jarak antar kedua obyek tersebut. Jarak paling dekat antara satu data dengan satu *cluster* tertentu akan menentukan suatu data masuk ke dalam *cluster* yang mana.
4. Hitung nilai centroid atau pusat *cluster* pada tiap-tiap *cluster*. Pusat *cluster* adalah rata-rata semua data atau obyek dalam *cluster*.
5. Alokasikan lagi setiap obyek memakai pusat *cluster* yang baru. Jika pusat *cluster* sudah tidak berubah lagi, maka proses peng-*cluster*-an selesai.
6. Kembali ke langkah 3 sampai pusat *cluster* tidak berubah lagi

3. Metode *K-Medoids*

K-Medoids atau *Partitioning Around Medoids (PAM)* adalah bagian dari *partitioning clustering* (Zayuka dkk, 2017) dan merupakan algoritma *clustering* yang mirip dengan *K-Means* (Pramesti dkk, 2017). Metode *K-Medoids* menggunakan objek perwakilan (*medoid*) sebagai pusat *cluster* atau titik acuan pada setiap *cluster*, bukan menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster* (Ramadhani dan Januarita, 2017). Algoritma *K-Medoids* ini mengambil parameter input *k* yang merepresentasikan jumlah *cluster* yang akan dipartisi diantara satu set sejumlah (*n*) objek (Ramadhani dan Januarita, 2017).

Langkah-langkah algoritma *K-Medoids* (Ramadhani dan Januarita, 2017), (Zayuka dkk, 2017):

1. Inisialisasi pusat *cluster* sebanyak *k* (jumlah *cluster*)

2. Alokasikan setiap data atau objek ke *cluster* terdekat dengan menggunakan *Euclidian Distance*, dengan persamaan :

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

dimana :

$d(x,y)$ = jarak antara data ke-*i* dan data ke-*j*

x_i = nilai atribut kesatu dari data ke-*i*

y_i = nilai atribut kesatu dari data ke-*j*

n = jumlah atribut yang digunakan

3. Pilih secara acak objek pada masing-masing *cluster* sebagai kandidat *medoid* baru
4. Hitung jarak setiap objek yang ada pada masing-masing *cluster* dengan kandidat *medoid* baru.
5. Hitung total simpangan (*S*) dengan menghitung nilai total *distance* yang baru dikurangi total *distance* lama (TD baru - TD lama). Jika $S < 0$, maka tukar objek dengan data *cluster* untuk membentuk sekumpulan *k* objek baru sebagai *medoid*.
6. Ulangi langkah 3 - 5 hingga tidak terjadi perubahan *medoid*.

4. Customer Relationship Management (CRM) dan Pelanggan Potensial

Customer Relationship Management (CRM) merefleksikan peranan utama dari konsumen untuk pengaturan strategi perusahaan yang meliputi seluruh ukuran untuk memahami konsumen dimana eksploitasi pengetahuan ini digunakan untuk merancang dan mengimplementasikan pada kegiatan marketing, produksi dan rantai *supply* dari pemasok (*supplier*). Beberapa definisi tentang CRM dijelaskan sebagai berikut (Kristanto dan Arief, 2013) :

- a. CRM dari sisi yang berkaitan dengan teknologi informasi dijelaskan sebagai sebuah cara atau strategi yang digunakan untuk mengoptimalkan customer lifetime value dengan cara mengetahui lebih banyak tentang informasi konsumen juga

- berinteraksi dengan konsumen secara intensif.
- b. CRM dari sisi komunikasi dan manajemen didefinisikan sebagai sebuah pendekatan perusahaan untuk bisa memahami juga mempengaruhi perilaku konsumen melalui komunikasi yang intensif untuk meningkatkan akuisisi konsumen, mempertahankan konsumen dan loyalitas konsumen.
 - c. CRM dari segi bisnis diartikan sebagai sebuah strategi bisnis untuk memahami, mengantisipasi dan mengelola kebutuhan konsumen yang potensial dalam suatu organisasi pada saat sekarang dan yang akan datang.

Menurut Zulkarnain (2012), pelanggan potensial adalah pelanggan yang loyal terhadap perusahaan. Pelanggan potensial biasanya akan melanjutkan pembelian produk dan jasa yang ditawarkan *seller* atau perusahaan walaupun dihadapkan pada banyak alternatif produk atau jasa yang lain yang lebih unggul dipandang dari berbagai sudut atribut yang ada pada produk atau jasa tersebut. Menurut Kotler dalam (Zulkarnain, 2012) konsumen atau pelanggan loyal tidak diukur dari berapa banyak orang tersebut membeli, tetapi dari seberapa sering dia melakukan transaksi pembelian ulang, termasuk merekomendasikan kepada orang lain untuk membeli.

5. Model Fuzzy RFM (Recency, Frequency dan Monetary)

Model RFM (*Recency, Frequency dan Monetary*) digunakan untuk menentukan variabel mengukur pembelian produk oleh pelanggan. Variabel dapat menentukan sebagai kebaruan, frekuensi dan moneter (Yuliaridkk, 2015).

- a. *Recency* adalah rentang waktu (hari, bulan, tahun) dari transaksi akhir sampai saat ini oleh pelanggan membeli.

- b. *Frequency* atau frekuensi adalah total transaksi atau transaksi rata-rata dalam satu kali periode.
- c. *Monetary* atau moneter adalah biaya rata-rata total pelanggan di sekali waktu.

C. METODE PENELITIAN

Populasi target dalam penelitian ini adalah data historis transaksi penjualan bulan Januari s/d Juni 2015 dengan data master pelanggan yang melakukan transaksi secara *online* saja. Data yang digunakan di uji dengan SPSS untuk mengetahui apakah data layak diproses selanjutnya atau tidak dengan menggunakan Uji Kelayakan *Kaiser Mayer Olkin* (KMO) dan *Barlett Test*. Jika nilai KMO kurang dari 0,5 maka analisis faktor tidak layak dilakukan. Pada uji kelayaka data yang digunakan di peroleh hasil nilai nilai KMO sebesar 0,590 sehingga data layak digunakan atau diproses lebih lanjut. Tampilan hasil uji KMO dapat dilihat dibawah ini :

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.590
Bartlett's Test of Sphericity	Approx. Chi-Square	20.769
	df	3
	Sig.	.000

Gambar III. 1 Nilai KMO

Metode Analisis pada penelitian ini menggunakan metode KDD yang terdiri dari sembilan langkah (Maimon dan Rokach, 2010), (Muningsih, 2015) yaitu :

1. Tahap *Domain Understanding* dan *KDD Goals*. Berdasarkan pengamatan, penentuan pelanggan potensial dilakukan secara manual sehingga tidak akurat.
2. Tahap *Selection* dan *Addition*. Data historis diambil dari data transaksi penjualan (*online*). Data transaksi penjualan (*online*) ditampilkan dibawah ini :

	Transaksi Terakhir	R	F	M	1		2		3		4			
					Tanggal	Jumlah	Tanggal	Jumlah	Tanggal	Jumlah	Tanggal	Jumlah	Tanggal	
D10002	1 Keranjang Batik Diana Jogja	5/11/2015	2	7	1.235.000	5/6/2015	260.000	3/20/2015	80.000	4/8/2015	210.000	4/14/2015	315.000	4/23/2015
D10005	2 Heni Olshop	6/23/2015	3	7	1.785.000	2/24/2015	350.000	3/26/2015	210.000	3/31/2015	190.000	5/3/2015	145.000	4/7/2015
D10009	3 Charen Shop (Efuida)	6/26/2015	3	10	925.000	3/23/2015	70.000	4/30/2015	95.000	4/22/2015	70.000	5/8/2015	95.000	5/21/2015
D10010	4 Evarista (Serafim Collection)	4/24/2015	1	2	214.000	4/10/2015	136.000	4/24/2015	78.000					
D10011	5 OLSHOP Syahabsyl	5/4/2015	2	2	230.000	4/27/2015	165.000	5/4/2015	65.000					
D10013	6 Siti Aisyah	6/29/2015	3	2	380.000	6/3/2015	285.000	6/29/2015	95.000					
C100003	7 Norliani	3/30/2015	1	2	155.000	3/4/2015	105.000	3/30/2015	50.000					
C100004	8 Denny Eka Anggraini	5/25/2015	2	5	859.000	3/4/2015	145.000	3/12/2015	420.000	3/28/2015	70.000	5/14/2015	165.000	5/25/2015
C100009	9 Lia	3/30/2015	1	4	703.000	3/9/2015	290.000	3/11/2015	130.000	3/12/2015	220.000	3/30/2015	63.000	
C100010	10 Nova PM Sunda	5/18/2015	2	2	275.000	3/7/2015	145.000	5/18/2015	130.000					
C100011	11 Icha Cahyawati	5/8/2015	2	2	1.970.000	3/7/2015	385.000	5/8/2015	1.585.000					
C100012	12 Susi	4/21/2015	1	3	493.000	3/7/2015	195.000	3/13/2015	103.000	4/21/2015	195.000			
C100013	13 Amriah	6/11/2015	3	4	1.544.000	3/9/2015	145.000	3/23/2015	670.000	5/9/2015	544.000	6/11/2015	185.000	
C100015	14 Lia Okta	6/22/2015	3	3	1.078.000	3/10/2015	103.000	4/22/2015	665.000	6/22/2015	310.000			
C100016	15 Yanti Budiharto	6/23/2015	3	2	540.000	6/9/2015	190.000	6/23/2015	350.000					
C100017	16 Rulany	6/3/2015	3	8	696.000	3/11/2015	70.000	3/23/2015	90.000	3/24/2015	90.000	3/28/2015	80.000	3/31/2015
C100018	17 Rina Magelang	3/12/2015	1	2	350.000	1/25/2015	40.000	3/12/2015	310.000					
C100019	18 Indah Ocktavia	4/27/2015	1	4	1.104.000	3/12/2015	372.000	3/25/2015	333.000	4/4/2015	216.000	4/27/2015	183.000	
C100024	19 Fitri Cahya Sari	6/24/2015	3	2	745.000	1/24/2015	185.000	6/24/2015	560.000					
C100027	20 Rahmah Mukmin	6/16/2015	3	2	575.000	1/30/2015	325.000	6/16/2015	250.000					
C100028	21 Ditha Angelina Makalagal	3/27/2015	1	2	490.000	1/25/2015	175.000	3/27/2015	315.000					
C100033	22 Fransiska Kiki	5/26/2015	2	2	720.000	4/20/2015	530.000	5/26/2015	190.000					
C100039	23 Tri Handayani	6/6/2015	3	1	700.000	6/6/2015	700.000							
C100052	24 Yuyun Prasetyo Ningrum	5/15/2015	2	12	1.990.900	3/20/2015	63.000	3/23/2015	63.000	3/25/2015	117.000	3/27/2015	86.400	3/27/2015
C100053	25 Ricka Hartati	4/17/2015	1	3	340.000	3/30/2015	160.000	4/3/2015	100.000	4/17/2015	80.000			
C100057	26 Artha Donda Martina	3/23/2015	1	2	96.000	3/18/2015	48.000	3/23/2015	48.000					

Gambar III. 2 Data Transaksi Penjualan (Online)

3. Tahap *Preprocessing* dan *Data Cleansing*. *Preprocessing* data pada tahap ini adalah mengambil data transaksi pelanggan yang melakukan transaksi minimal 2 kali dan didapatkan pelanggan sebanyak 42 pelanggan
4. Tahap *Transformation*. Proses transformasi data dengan cara memberikan kode pelanggan kepada pelanggan.
5. Tahap Data Mining memilih tipe *data mining* yang cocok. Tipe *Data Mining* yang digunakan dalam penelitian ini adalah *clustering* untuk pengelompokan pelanggan potensial
6. Tahap Data Mining memilih algoritma *data mining*. Metode yang digunakan adalah metode *clustering K-Means* dan metode *K-Medoids* dengan model *fuzzyRFM* dimana
 7. parameter yang digunakan adalah variabel *Recency* yang menyatakan atribut Tanggal Transaksi Terakhir, *Frequency* menyatakan Jumlah Transaksi, dan *Monetary* menyatakan atribut Total Uang yang dibelanjakan yang masing-masing dibagi menjadi 3 kelompok *fuzzy* yaitu :
 - a. *Recency* dibagi menjadi kelompok *fuzzy* Baru Saja dengan nilai 3, Agak Lama dengan nilai 2 dan Lama dengan nilai 1.
 - b. *Frequency* dibagi menjadi kelompok *fuzzy* Sering dengan nilai 3, Agak Sering dengan nilai 2 dan Jarang dengan nilai 1.
 - c. *Monetary* dibagi menjadi kelompok *fuzzy* Tinggi dengan nilai 3, Sedang dengan nilai 2 dan Rendah dengan nilai 1.

Tabel III. 1 Nilai Parameter RFM

No	Parameter	nilai	<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>
1.	Besar		$0 < r < 31$ hr	$0 < f < 3$	$m \geq 1.500.000$
2.	Sedang		$31 \text{ hr} < r < 60$ hr	$3 < f < 6$	$500.000 < m < 1.500.000$
3.	Rendah		$r \geq 61$ hr	$f \geq 6$	$0 < m < 500.000$

8. Tahap *Data Mining* dalam penggunaan algoritma *data mining*. Pada tahap ini dilakukan implementasi dari algoritma *data mining* yang telah ditentukan pada tahap sebelumnya.

9. Tahap *Evaluation*. Pada tahap ini dilakukan evaluasi terhadap sampel data yang dihasilkan mengenai penentuan *clustering* pelanggan potensial. Untuk selanjutnya proses analisis data, akan dihitung tingkat akurasi dengan menggunakan penghitungan metode *Fuzzy RFM* (*recency*, *frequency* dan *monetary*)
10. Tahap *Discovered Knowledge* (*Visualization* dan *Integration*). Penggunaan pengetahuan yang diperoleh. Pada tahap ini, penerapan metode *clustering* terbaik atau paling akurat dengan model *Fuzzy RFM* untuk *clustering* penentuan kelompok pelanggan potensial.

D. PEMBAHASAN

1. Pengolahan Data

Pengolahan data yang dilakukan pada penelitian ini menggunakan aplikasi *Rapid Miner 9.0*. Dari pengolahan data yang sudah dilakukan dengan menggunakan jumlah *cluster* 3 didapatkan hasil untuk jumlah anggota dan *centroid* (titik pusat) masing-masing *cluster* pada metode *K-Means* dan metode *K-Medoids* yang ditampilkan pada tabel dibawah ini :

Tabel IV.1 Jumlah Anggota *Cluster*

Metode	<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>
<i>K-Means</i>	16	11	15
<i>K-Medoids</i>	13	9	20

Tabel IV.2 *Centroid* Metode *K-Means*

Atribut	<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>
<i>R</i>	1,188	2,818	2,733
<i>F</i>	1,312	1,091	2,400
<i>M</i>	1,250	1,909	2,333

Tabel IV.3 *Centroid* Metode *K-Medoids*

Atribut	<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>
---------	------------------	------------------	------------------

<i>R</i>	1	2	3
<i>F</i>	1	2	2
<i>M</i>	1	2	2

2. Pengujian Akurasi

Cluster yang terbentuk berdasarkan faktor *recency*(*R*), *frequency*(*F*) dan *monetary*(*M*). Semakin besar nilai *R* menunjukkan bahwa pelanggan sering melakukan transaksi, semakin besar nilai *F* menunjukkan pelanggan setia terhadap terhadap produk yang digunakan, dan semakin besar nilai *M*, menunjukkan bahwa nominal transaksi yang dibayarkan semakin besar. Pengujian dilakukan dengan mengambil rata-rata (mean) dari nilai *Fuzzy RFM*(*recency*, *frequency* dan *monetary*), dimana jika nilai rata-rata lebih kecil sama dengan 1,3 maka merupakan anggota *cluster 1*, jika nilai rata-rata lebih besar dari 1,3 dan lebih kecil dari 2,3 merupakan anggota *cluster 2*, dan jika nilai rata-rata lebih besar sama dengan 2,3 maka merupakan anggota *cluster 3*. Pengujian akurasi metode *K-Means* dan *K-Medoids* ditampilkan pada tabel dibawah ini :

Tabel IV.4 Pengujian Akurasi Metode *Clustering*

Kode Plg	Pengujian	<i>K-Means</i>	<i>K-Medoids</i>
P1	<i>Cluster 3</i>	akurat	tidak akurat
P2	<i>Cluster 3</i>	akurat	akurat
P3	<i>Cluster 3</i>	akurat	akurat
P4	<i>Cluster 1</i>	akurat	akurat
P5	<i>Cluster 1</i>	akurat	akurat
P6	<i>Cluster 2</i>	akurat	tidak akurat
P7	<i>Cluster 1</i>	akurat	akurat
P8	<i>Cluster 2</i>	tidak akurat	akurat
P9	<i>Cluster 2</i>	akurat	akurat
P10	<i>Cluster 1</i>	akurat	akurat

P11	Cluster 2	akurat	akurat
P12	Cluster 1	akurat	akurat
P13	Cluster 3	akurat	akurat
P14	Cluster 3	akurat	akurat
P15	Cluster 2	akurat	tidak akurat
.....
P42	Cluster 1	akurat	akurat

Hasil dari pengujian akurasi yang dilakukan diperoleh hasil yaitu metode *K-Means* dengan nilai akurat sebanyak 37 dan tidak akurat sebanyak 4 orang. Sedangkan metode *K-Medoids* memiliki nilai akurasi 32 orang dan tidak akurat sebanyak 10 dari jumlah data 42. Sehingga jika diprosentasekan nilai akurasi metode *K-Means* adalah 90,47% dan metode *K-Medoids* 76,19%. Sehingga didapatkan kesimpulan bahwa metode *K-Means* lebih akurat dan lebih baik dibanding metode *K-Medoids*. Dan hasil pengklasteran dengan metode *K-Means* menghasilkan 3 (tiga) kelompok pelanggan yaitu :

1. Cluster 1 : 16 anggota pelanggan kode P4, P5, P7, P9, P10, P12, P17, P18, P21, P25, P26, P29, P31, P34, P41 dan P42 dengan pusat cluster (1,188; 1,312; 1,250), memiliki nilai rata-rata ketiga faktor yang kecil, sehingga dapat digolongkan menjadi pelanggan yang tidak atau kurang potensial.
2. Cluster 2 : 11 anggota pelanggan kode P6, P11, P15, P19, P20, P22, P23, P27, P32, P37 dan P38 dengan pusat cluster (2,818; 1,091; 1,909), memiliki nilai rata-rata ketiga nilai yang cukup besar, sehingga dapat digolongkan sebagai pelanggan yang potensial.
3. Cluster 3 : 15 anggota pelanggan P1, P2, P3, P8, P13, P14, P16, P24, P28, P30, P33, P35, P36, P39 dan P40 dengan pusat cluster (2,733; 2,400; 2,333), memiliki nilai yang besar dari ketiga faktor, sehingga dapat digolongkan sebagai pelanggan yang sangat potensial.

4. Kesimpulan

Dari penelitian yang dilakukan, metode *K-Means* dan *K-Medoids* dengan model *Fuzzy RFM* terbukti bahwa metode *K-Means* lebih baik dari metode *K-Medoids*. Metode *K-Means* dengan akurasi 90,47% dapat melakukan pengelompokan pelanggan menjadi 3 kluster yaitu pelanggan kurang atau tidak potensial, pelanggan potensial dan pelanggan sangat potensial. Dengan memperlakukan pelanggan sesuai karakteristiknya, diharapkan hasil penelitian diharapkan dapat memaksimalkan strategi pemasaran. Walaupun hasil yang diharapkan sudah terpenuhi, namun untuk penelitian lebih lanjut bisa dikembangkan dengan menggunakan metode *Data Mining* yang lain.

DAFTAR PUSTAKA

Agusta, Y., 2007, *K-Means - Penerapan, Permasalahan dan Metode Terkait*. *Jurnal Sistem dan Informatika*, vol 3, hal 47-60.

Febrianti, K., Hafiyusholeh, Moh., Asyuhar, A.H., 2016, *Perbandingan Pengklasteran Data Iris Menggunakan Metode K-Means dan Fuzzy C-Means*, *Jurnal Matematika "MANTIK"*, vol 2 no 01, hal 7-13

Kristanto, T., & Arief, R., 2013, *Analisa Data Mining Metode Fuzzy Untuk Customer Relationship Management Pada Perusahaan Tour & Travel*. *Seminar Nasional Sistem Informasi Indonesia* (hal. 528-534). Surabaya: Sesindo ITS.

Larose, D. T., 2005. *Discovering Knowledge in Data : An Introduction to Data Mining*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Maimon, O., & Rokach, L., 2010, *Data Mining and Knowledge Discovery Handbook*. London: Springer Science+Business Media.

Moertini, V. S., 2002. *Data Mining Sebagai Solusi Bisnis*. *Integral*, Vol 7 (No. 1).

Muningsih, E., 2015, *Penerapan Metode Fuzzy C-Means Dengan Model Fuzzy RFM (Studi Kasus : Clustering Pelanggan Potensial Online Shop*. *Proceeding Seminar Nasional Teknologi Informasi dan Komunikasi Terapan (SEMANTIK) 2015*, Universitas Dian Nuswantoro (UDINUS) Semarang, November 21.

Pramesti, D.F., Furqon, M.Tanzil., and Dewi, C., 2017, *Implementasi Metode K-Medoids*

Clustering untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hotspot), *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol 1 no 9, hal 723-732

Ramadhani, R.D., Januarita, Dwi AK., 2017, Evaluasi K-Means dan K-Medoids pada Dasaset Kecil, *Seminar Nasional Informatika dan Aplikasinya (SNIA) 2017*, Cimahi, September 2017.

Susanto, S., & Suryadi, D., 2010, *Pengantar Data Mining : Menggali Pengetahuan dari Bongkahan Data*. Yogyakarta: Penerbit ANDI.

Yuliani, N.P.P., Putra, I.K.G.D., dan Rusjayanti. N.K.D., 2015, Customer Segmentation Through Fuzzy C-Means And Fuzzy RFM Method". *Journal of Theoretical and Applied Information Technology*, Vol. 78, No. 3, 380-385

Zayuka, H., Nasution, S, M., and Purwanto, Y., 2017, Perancangan dan Analisis Clustering Data Menggunakan Metode K-Medoids untuk Berita Berbahasa Inggris, *e-Proceeding of Engineering Vol 4. No 2*, Agust 2

Zulkarnain, 2012, *Ilmu Menjual*. Yogyakarta : Graha Ilmu.