ANALISIS PERFORMA ALGORITMA *FUZZY C-MEANS* DAN *K-MEANS CLUSTERING* UNTUK PENGELOMPOKAN PELANGGAN PADA PT. PART STATION JEMBER

[1] Toni Hendrawan Nata Utama (15 1065 1068), [2] Ilham Saifudin, S.Pd., M.si, [3]Ari Eko Wardoyo, M. Kom.

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember
Jl. Karimata No. 49, Kotak Pos 104 Jember, 68121, Indonesia
Email: toni.h.n.u@gmail.com

ABSTRAK

Penelitian ini bertujuan menganalisis hasil performa algoritma Fuzzy C-Means dan K-Means Clustering untuk pengelompokan pelanggan pada perusahaan Part Station Jember. Dibedakan berdasarkan transaksi kredit paling rendah dari semua Cluster, transaksi tunai tertinggi dari semua Cluster, dan laba yang dihasilkan tertinggi dari semua Cluster. Dibentuk menggunakan metode uji validitas kelompok Silhouette Coefficient dan RMSSTD (Root Mean Square Standard Deviation). Kelompok optimum yang dihasilkan dari metode elbow ada 3.

Hasil pengujian metode Silhouette Coefficient pada data kredit menunjukkan algoritma Fuzzy C-Means lebih baik dibandingkan algoritma K-Means Clustering dengan nilai 0.61 untuk algoritma Fuzzy C-Means dan 0.59 untuk algoritma K-Means Clustering. Sedangkan pada pengujian data tunai dan laba menunjukkan algoritma K-Means Clustering lebih baik dibandingkan Fuzzy C-Means dengan nilai masing-masing 0.66 dan 0.25 untuk algoritma Fuzzy C-Means dan 0.79 dan 0.4 untuk algoritma K-Means Clustering. hal ini menunjukkan metode validasi Silhouette Coefficient ini sangat dipengaruhi oleh data yang digunakan. Akan tetapi hasil pengujian metode RMSSTD menunjukkan algoritma Fuzzy C-Means lebih baik dibandingkan algoritma K-Means Clustering dari semua data yang dilakukan pengujian, baik itu data kredit, tunai dan laba. Dengan nilai rata-rata 0.024979667 data kredit, 0.06873552 data tunai, 0.010187823 data laba untuk algoritma Fuzzy C-Means sedangkan nilai rata-rata 0.03229311 data kredit, 0.105352232 data tunai, 0.064485279 data laba untuk algoritma K-Means Clustering.

Kata Kunci : K-Means Clustering, Fuzzy C-Means, Silhouette Coefficient, RMSSTD

ABSTRACT

This study aims to analyze the results of the performance of the Fuzzy C-Means and K-Means Clustering algorithms for grouping customers in the Jember Part Station company. Differentiated based on the lowest credit transactions of all Clusters, the highest cash transactions of all Clusters, and the highest profit generated from all Clusters. Formed using the method of testing the validity of the Silhouette Coefficient group and RMSSTD (Root Mean Square Standard Deviation). The optimum group produced from the elbow method is 3.

The results of testing the Silhouette Coefficient method on credit data show that the Fuzzy C-Means algorithm is better than the K-Means Clustering algorithm with a value of 0.61 for the Fuzzy C-Means algorithm and 0.59 for the K-Means Clustering algorithm. While the testing of cash and profit data shows that the K-Means Clustering algorithm is better than Fuzzy C-Means with values of 0.66 and 0.25 for the Fuzzy C-Means algorithm and 0.79 and 0.4 for the K-Means Clustering algorithm, respectively. this shows that the Silhouette Coefficient validation method is strongly influenced by the data used. However, the results of testing the RMSSTD method show that the Fuzzy C-Means algorithm is better than the K-Means Clustering algorithm of all data tested, both credit, cash and profit data. With an average value of 0.024979667 credit data, 0.06873552 cash data, 0.010187823 profit data for the Fuzzy C-Means algorithm while the average value is 0.03229311 credit data, 0.105352232 cash data, 0.064485279 profit data for the K-Means Clustering algorithm.

***Keywords**:K-Means Clustering, Fuzzy C-Means, Silhouette Coefficient, RMSSTD*

PENDAHULUAN

Dalam dunia bisnis terutama pada sektor perdagangan peran pelanggan sangat penting untuk meningkatkan efektifitas pemasaran dan mengoptimalkan keuntungan. Pelaku bisnis harus mampu untuk mengenali kelompok pelanggan dan mempercayainya dengan meningkatkan pemahaman perusahaan akan kebutuhan mereka sehingga dapat mempertahankan loyalitasnya terhadap perusahaan. untuk bisa mengetahui hal – hal diatas maka terdapat cara yaitu dengan melakukan analisis data perusahaan, dalam hal ini analisis data dilakukan pada perusahaan Part Station Jember yang bergerak di bidang distributor spare part atau suku cadang sepeda motor.

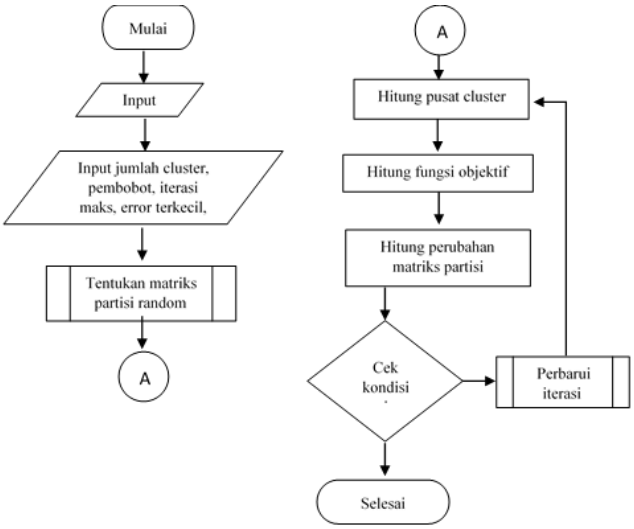
Data pada perusahaan *Part Station* Jember dapat diekstraksi dan dilakukan identifikasi informasi – informasi baru yang berguna untuk meningkatkan efektivitas bisnis perusahaan. Data – data tersebut digali kemudian diolah menjadi sebuah informasi yang bermanfaat dengan sebuah metode disebut *Data Mining* (Dawson, 2009).

Ketersediaan data yang cukup banyak berupa data transaksi nota, serta kebutuhan akan informasi atau pengetahuan untuk menunjang kinerja perusahaan. Data transaksi berupa nota yang di dalamnya meliputi transaksi kredit, transaksi tunai, dan laba akan dianalisis untuk pengelompokan pelanggan yaitu pelanggan yang dikatakan memiliki kualitas bagus. Oleh karena itu dengan menerapkan metode algoritma *Fuzzy C-Means* dan *K-Means Clustering* dapat membantu menganalisa data untuk diproses sehingga menghasilkan pengetahuan baru untuk pengelompokan pelanggan di perusahaan distributor Part Station Jember.

TINJAUAN PUSTAKA

a. Fuzzy C-Means

Fuzzy C-Means pertama kali ditemukan oleh Bezdek pada tahun 1981. Fuzzy C-Means adalah teknik pengklasteran data dimana keberadaan pada setiap titik data dikelompokkan dalam suatu cluster dengan derajat keanggotaan tertentu. Pengelompokan data dengan FCM menghasilkan keluaran berupa daftar pusat cluster dan beberapa fungsi keanggotaan untuk mempresentasikan nilai fuzzy dari setiap cluster (Yohanes, 2016). Berbeda halnya dengan K-means, data yang di cluster menggunakan FCM akan menjadi anggota dari setiap cluster yang ada. Ikatan data dengan cluster ditentukan oleh nilai keanggotaannya yang berada pada rentan 0 hingga 1. Prosedur algoritma *Fuzzy C-Means* ditunjukkan pada Gambar 1

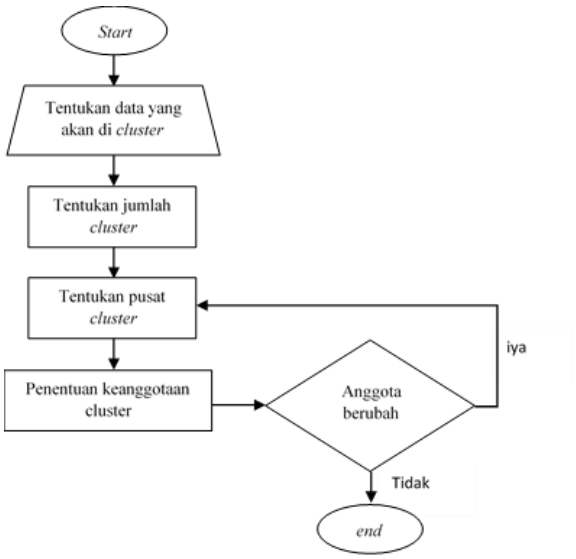


Gambar 1 Prosedur Algoritma *Fuzzy C-Means*

b. *K-Means Clustering*

K-Means Clustering ditemukan oleh MacQueen pada tahun 1967 (Yohanes, 2016). K- Means adalah salah satu teknik pengelompokan yang paling banyak digunakan karena kesederhanaan dan kecepatan. K-Means bekerja dengan mengelompokkan data yang ada ke dalam beberapa cluster. Data akan dikumpulkan dalam satu cluster berdasarkan kemiripan ataupun kedekatan karakteristik dengan data yang lain yang ada di cluster itu. K-Means akan berusaha menjauhkan data pada suatu cluster dengan cluster lainnya. Hal ini dilakukan untuk meminimalkan variasi antar data yang berada dalam satu cluster dan memaksimalkan variasi dengan data yang berada di cluster lainnya.

Algoritma K-Means Clustering adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (Unsupervised dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok , dimana data dalam satu kelompok mempunyai karakteristik yang sama satu dengan lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain (Agusta, 2011). rosedur algoritma *K-Means Clustering* Gambar 2.



Gambar 2 Prosedur Algoritma *K-Means Clustering*

c. *Silhouette Coefficient*

Silhouette Coefficient merupakan salah satu metode validasi untuk menguji kualitas sebuah cluster (Tanzil Furqon & Muflikhah, 2016). Metode ini menggunakan rumus Eculidean distance dalam proses perhitungannya. Secara singkat metode Silhouette Coefficient ditunjukkan pada persamaan berikut:

$$s_i = \frac{(b_i - a_i)}{\max(b_i\ ,\ a_i)}$$

s_i = *Silhouette Coefficient*

a_i = rata – rata jarak dari objek i dengan seluruh objek yang berada pada *cluster* yang sama

b_i = nilai terkecil dari rata – rata jarak objek i dengan objek lain pada cluster yang berbeda

Nilai Silhouette Coefficient berada pada rentan (-1) hingga 1. Semakin tinggi nilainya, maka semakin bagus pula kualitasnya.

d. RMSSTD *(Root Mean Square Standard Deviation)*

Root Mean Square Standard Deviation merupakan metode uji validitas kelompok untuk mengevaluasi hasil dari analisis kelompok secara kuantitatif sehingga dihasilkan kelompok optimum. Kelompok optimum adalah kelompok yang mempunyai jarak yang padat antar individu dalam kelompok dan terisolasi dari kelompok lain dengan baik. Menurut Sharma, (1996:197) indeks yang bisa dijadikan tolok ukur dalam pengujian validitas kelompok adalah indeks RMSSTD (Root Mean Square Standard Deviation).

$$d_{(UV)W} = d\sqrt{\frac{\sum_{i=1}^{n_c}\sum_{j=1}^{r_i}(x_{ij} - \underline{x_i})^2}{\sum_{i=1}^{n_c}(r_i - 1)}}$$

x_{ij}= nilai objek ke$-j$ pada kelompok i

$\overline{x_i}$ = nilai pusat kelompok ke-i

n_c= banyaknya kelompok yang terbentuk

r_i = banyaknya objek yang termasuk dalam kelompok i

RMSSTD mengukur kehomogenan dari kelompok yang terbentuk pada setiap tahap. Semakin kecil nilai RMSSTD semakin tinggi nilai kehomogenan kelompok yang terbentuk pada tahap tersebut. Nilai RMSSTD yang besar menunjukkan kluster yang terbentuk tidak homogen (Yatskiv, 2004).

e. Metode Elbow

Menurut Madhulatha (2012) dalam Merliana, et al (2015) disebutkan bahwa metode Elbow merupakan suatu metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah cluster optimum dengan cara melihat persentase hasil perbandingan antara jumlah cluster yang akan membentuk siku pada suatu titik. Metode elbow diberntuk dari persamaan SSE (*sum of square error)* seperti dibawah ini:

$$\text{SSE} = \sum_{k-1}^{k}\sum_{xi\in Sk}||Xi - Ck||_2^2$$

Xi = fitur atau atribut dari data ke-i

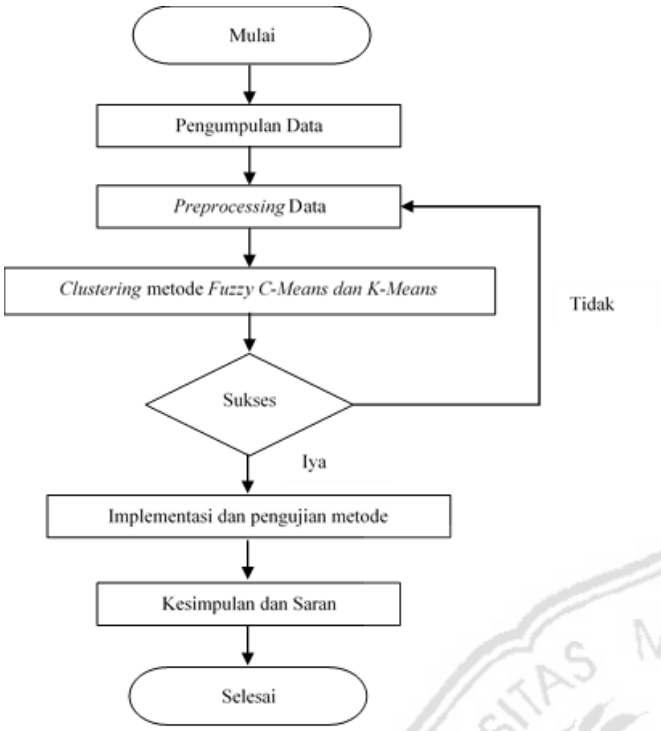
Ck = fitur atau atribut titik pusat cluster ke-i

METODOLOGI PENELITIAN

a. Jenis Penelitian

Penjelasan mengenai tahapan-tahapan dalam performa algoritma Fuzzy C-Means dengan K-Means pada pengelompokan pelanggan pada Part Station Jember. Metodologi penelitian yang dilakukan dalam penelitian ini melalui beberapa tahapan yaitu studi literatur, pengumpulan data, pra proses data, proses clustering, performa metode Fuzzy C-Means dan K-Means Clustering, dan pengambilan kesimpulan.

Diagram blog metodologi penelitian yang berisi tahapan-tahapan yang dilakukan dalam penelitian di tunjukkan pada Gambar 3.



Gambar 3 Diagram blog metodologi penelitian

b. Preses *Clustering*

Adapun proses clustering algoritma *Fuzzy C-Means* adalah sebagai berikut:

1. Tentukan jumlah *cluster*, maksimum iterasi, dan error terkecil.
2. Menentukan jumlah frekuensi pusat cluster awal, matrik pusat cluster ditentukan dengan persamaan sebagai berikut:

$$V_{kj} = \frac{\sum_{i=1}^{25}((\mu_{ik})^2 * X_{ij})}{\sum_{i=1}^{25}(\mu_{ik})^2}$$

Dimana:

k = *cluster*

i = iterasi (1)

μ_{ik} = perubahan matrik partisi

X_{ij} = Atribut transaksi nota

3. Menghitung fungsi objektif (Po), fungsi objektif dapat dihitung menggunakan persamaan berikut:

$$P_{1} = \sum_{i=1}^{25}\sum_{k=1}^{3}\left(\left[\sum_{j=1}^{3}(X_{ij} - V_{kj})^2\right](\mu_{ik})^2\right)$$

Dimana:

p_1 = Fungsi objektif

$\sum_{i=1}^{25}$ = Jumlah data yang di *cluster*

$\sum_{k=1}^{3}$ =Jumlah perhitungan *cluster* awal

4. Menghitung perubahan nilai keanggotaan cluster awal yang dihitung menggunakan persamaan:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{4}(X_{ij} - V_{kj})^2\right]^{\frac{-1}{2-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{4}(X_{ij} - V_{kj})^2\right]^{\frac{-1}{2-1}}}$$

5. Melakukan perulangan tahap 9 sampai 11 hingga memenuhi syarat sebagai berikut:
   - Jika = (|Pt-Pt-1|<$\varepsilon_{\text{ı}}$)
   - Jika = t>Maxlter

   Dimana:

   |Pt-Pt-1| = Fungsi objektif awal dikurangi fungsi objektif baru

   $\varepsilon_{\text{ı}}$ = Error terkecil

   t = Iterasi awal

   Maxter = Maksimum iterasi

Adapun proses clustering algoritma *K-Means Clustering* adalah sebagai berikut:

1. Tentukan jumlah *cluster*, maksimum iterasi, dan pusat cluster awal.
2. Menghitung jarak setiap data transaksi yang telah diinputkan terhadap masing-masing centroid menggunakan rumus jarak Eucledian Distance hingga ditemukan jarak paling dekat dari setiap data dengan centroid.

   Berikut adalah persamaan *Eucledian Distance:*

$$\sqrt{(x_{ia} - \mu_{ja})^2 + (x_{ia} - \mu_{ja})^2} \ldots\ldots 12$$

Dimana:

xi = data transaksi nota

μj = *centroid* pada *cluster* ke-j

xia = transaksi, tunai, kredit, laba

μja = nilai kriteria 1 dari *centroid cluster* ke-j adalah nilai random

3. Mengklasifikasikan setiap data transaksi berdasarkan kedekatan nya dengan *centroid* (jarak terkecil).
4. Memperbaharui nilai *centroid*. Nilai *centroid* baru diperoleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus:

$$\mu j(t+1) = \frac{1}{Nsj} \sum_{j \in Sj} x_j \quad \dots\dots\dots\dots$$

Dimana:

μj(t+1) = centroid baru pada iterasi ke (t+1)

Nsj = banyak data transaksi pada cluster Sj

5. Melakukan perulangan 2 sampai 5 hingga anggota dari centroid awal dan centroid baru tidak berubah.

IMPLEMENTASI DAN PENGUJIAN

a. Data Pengujian

Pada penelitian ini pengumpulan data dari PT. Part Station Jember dilakukan dengan pendekatan deskriptif. Selanjutnya data yang sudah dikumpulkan akan dilakukan preprocessing data tahap ini akan memilah atribut yang akan dipakai atau yang tidak berpengaruh pada pengujian dan pengolahan selanjutnya, atribut yang dipakai antara lain yaitu nama toko, id toko, transaksi tunai, transaksi kredit, dan laba. Data pengujian atau testing sendiri merupakan data yang sudah dilakukan preprocessing dan siap untuk dijadikan sebagai data uji, data uji atau testing berjumlah 240 data. Rincian dari data pengujian dapat dilihat pada tabel 1, 2, 3

No	Id Toko	Nama Toko	Kredit 1	…	Kredit 12
1	1031925	45 MOTOR	2444100	…	0
2	9906195	ABADI JAYA	1324000	…	260000
3	9906046	ABADI JAYA MOTOR	518600	…	0
4	9906493	ABEL MOTOR	238000	…	0
5	1058665	ADHITA MOTOR II	2391640	…	0
..	…	…	…	…	…
240	1053086	YURO MOTOR	683800	…	2968000

Tabel 1 Data *testing* kredit

No	Id Toko	Nama Toko	Kredit 1	…	Kredit 12
1	1031925	45 MOTOR	2444100	…	0
2	9906195	ABADI JAYA	1324000	…	260000
3	9906046	ABADI JAYA MOTOR	518600	…	0
4	9906493	ABEL MOTOR	238000	…	0
5	1058665	ADHITA MOTOR II	2391640	…	0
..	…	…	…	…	…
240	1053086	YURO MOTOR	683800	…	2968000

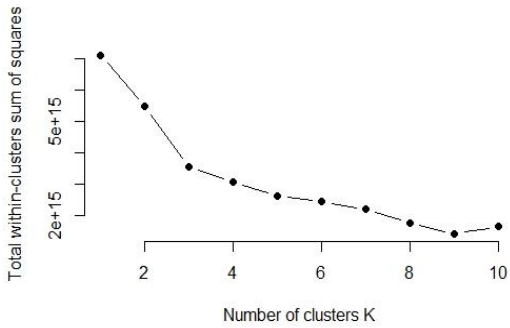
Tabel 2 Data *testing* tunai

No	Id Toko	Nama Toko	Kredit 1	…	Kredit 12
1	1031925	45 MOTOR	2444100	…	0
2	9906195	ABADI JAYA	1324000	…	260000
3	9906046	ABADI JAYA MOTOR	518600	…	0
4	9906493	ABEL MOTOR	238000	…	0
5	1058665	ADHITA MOTOR II	2391640	…	0
..	…	…	…	…	…
240	1053086	YURO MOTOR	683800	…	2968000

Tabel 3 Data *testing* laba

b. Mencari Optimum Cluster Menggunakan Metode Elbow

Pada tahap ini peneliti mencari jumlah cluster optimum untuk data yang akan dikelompokkan, penentuan nilai K dilakukan menggunakan metode Elbow. Hasil perhitungan metode Elbow ditampilkan dalam plot dua dimensi pada gambar 4 Nilai pada sumbu X adalah jumlah K yaitu 2 hingga 10, sedangkan sumbu Y adalah nilai SSE yang dihasilkan oleh setiap K. Metode Elbow menjelaskan

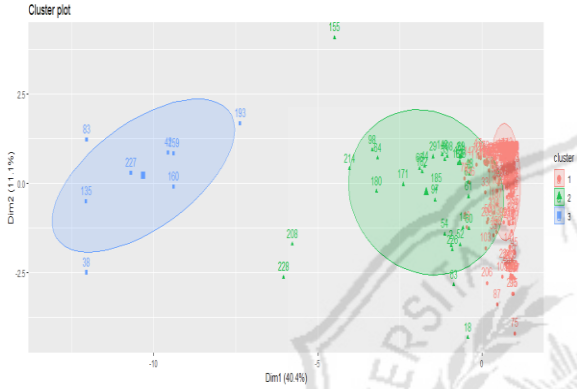


Gambar 4 Grafik Metode *Elbow*

bahwa nilai K yang diambil adalah pada titik dimana terjadi penurunan yang signifikan dan diikuti oleh nilai yang relatif konstan. Dapat dilihat bahwa titik yang dimaksud yaitu K=3.

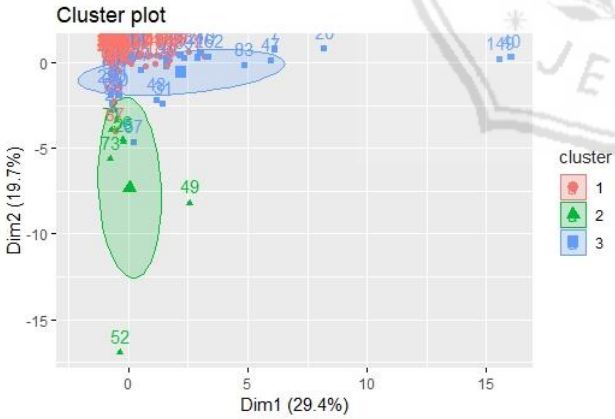
c. Distribusi Sebaran Data *Fuzzy C-Means* dan *K-Means*

Hasil dari sebaran cluster dan inisialisasi dari masing-anggota cluster data testing kredit, tunai, dan laba *Fuzzy C-Means* akan ditampilkan dalam bentuk plot yang diproses menggunakan RStudio seperti pada gambar 4, 5, 6, 7, 8, 9.



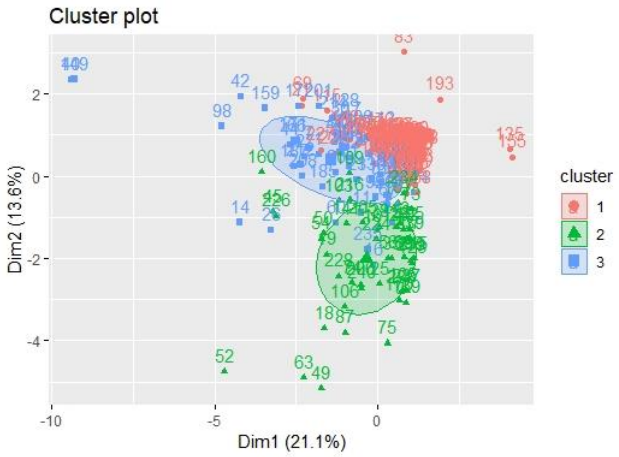

```
> res.fcm$csize
  1   2   3
199  33   8
```

Gambar 5 Tampilan Plot data Kredit *Fuzzy C-Means*



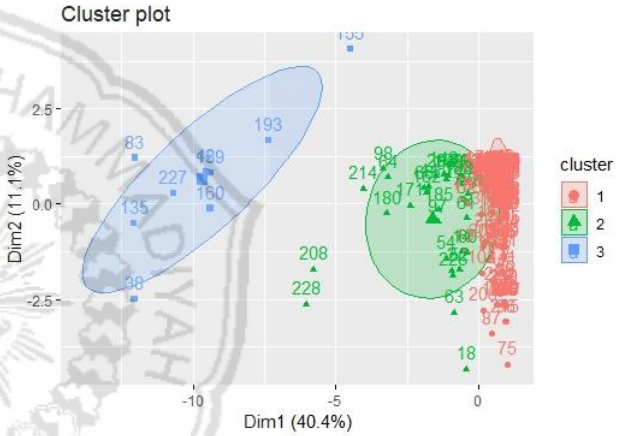

```
> res.fcm$csize
  1   2   3
202   6  32
```

Gambar 6 Tampilan Plot data Tunai *Fuzzy C-Means*



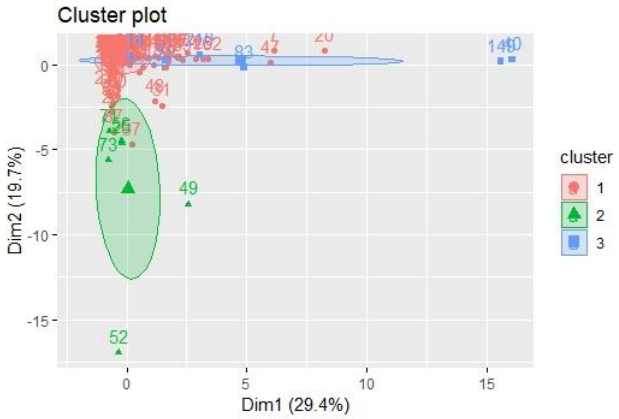

```
> res.fcm$csize
  1   2   3
134  48  58
```

Gambar 7 Tampilan Plot data Laba *Fuzzy C-Means*



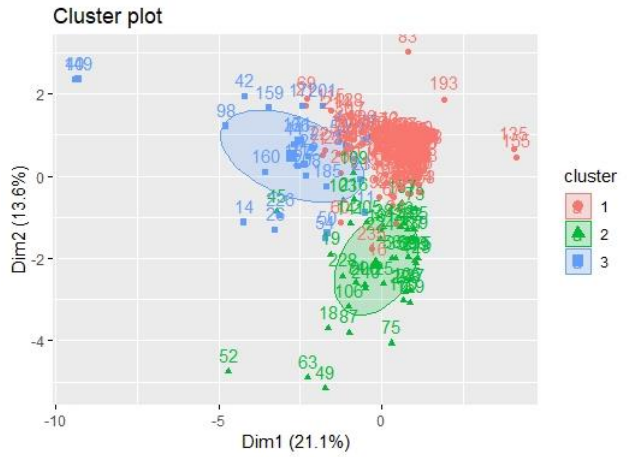

```
> res.kmeans$size
[1] 199  32   9
>
```

Gambar 8 Tampilan Plot data Kredit *K-Means*




```
> res.kmeans$size
[1] 225   6   9
>
```

Gambar 9 Tampilan Plot data Tunai *K-Means*




```
> res.kmeans$size
[1] 168  43  29
>
```

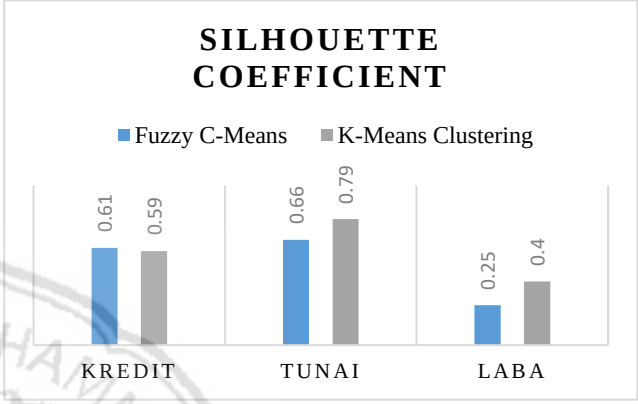
Gambar 10 Tampilan Plot data Laba *K-Means*

Dari gambar 4, 5, 6, 7, 8, 9, dapat dilihat bahwa kelompok yang terbentuk ada 3 yang berupa titik-titik dan nomor data, titik yang berwarna merah merupakan anggota dari cluster 1, titik hijau merupakan anggota cluster 2, dan titik biru merupakan anggota dari cluster 3. dimana dari plot tersebut dapat diketahui sebaran data pada masing-masing cluster dan jumlah anggotanya. Cluster yang dikatakan optimum yaitu yang memiliki jarak kecil pada tiap anggotanya dan terisolasi dari kelompok lain dengan baik. Anggota cluster terbentuk dari nilai keanggotaan pada iterasi terakhir.

- Pada data kredit, cluster 1 adalah kelompok data pelanggan dengan tingkat kredit rendah, cluster 2 adalah pelanggan dengan tingkat kredit sedang, dan cluster 3 merupakan pelanggan dengan tingkat kredit tinggi.
- Pada data tunai, cluster 1 adalah kelompok pelanggan dengan tingkat transaksi tunai rendah, cluster 2 merupakan pelanggan dengan tingkat transaksi tunai sedang, dan cluster 3 merupakan pelanggan dengan tingkat transaksi tunai tinggi.
- Pada data laba, cluster 1 merupakan kelompok pelanggan yang menghasilkan laba rendah, cluster 2 merupakan pelanggan yang menghasilkan laba sedang, dan cluster 3 merupakan pelanggan yang menghasilkan laba tinggi.

d. Penerapan Metode *Silhouette Coefficient* Pada *Fuzzy C-Means* dan *K-Means Clustering*

Pada proses penerapan metode Silhouette Coefficient ini, masing-masing data kredit, tunai, dan laba diuji satu persatu dan akan dilihat nilai yang terbaik yaitu data yang mempunyai nilai Silhouette Coefficient tinggi. Nilai Silhouette Coefficient berada pada rentan (-1) hingga 1. Semakin tinggi nilainya, maka semakin bagus pula kualitasnya.



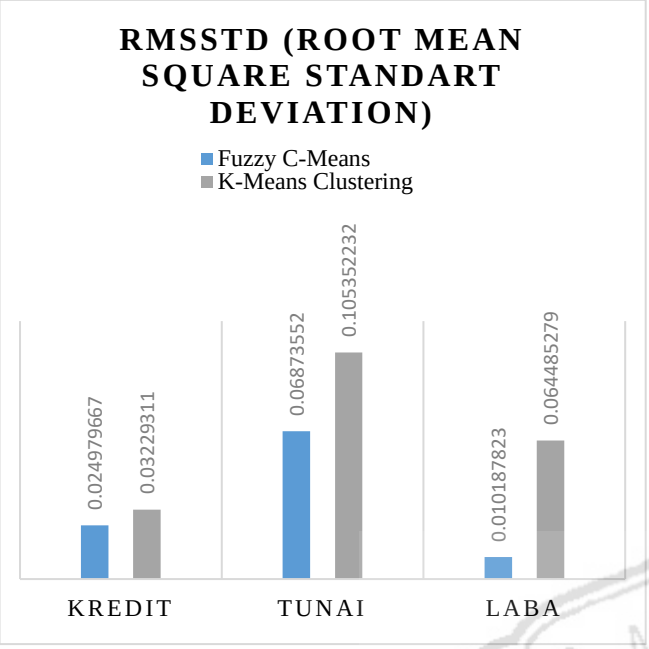
Gambar 11 Grafik hasil metode *Silhouette Coefficient*

Gambar 10 terlihat bahwa pada pengujian data kredit nilai Silhouette Coefficient algoritma Fuzzy C-Means lebih tinggi yaitu sebesar 0.61 dan algoritma K-Means sebesar 0.59, sedangkan pada pengujian data tunai dan laba algoritma K-Means clustering cenderung lebih tinggi secara signifikan dengan nilai 0.66 dan 0.25 untuk Fuzzy C-Means sedangkan 0.79 dan 0.4 untuk K-Means. Ketidak-selarasan hasil pengujian Silhouette Coefficient ini sangat dipengaruhi oleh data yang digunakan.

e. Penerapan Metode RMSSTD (*Root Mean Square Standard Deviation*) Pada *Fuzzy C-Means* dan *K-Means* Clustering

Konsep metode RMSSTD *(Root Mean Square Standard Deviation)* sama dengan metode *Silhouette Coefficient* yaitu untuk validasi hasil cluster. Metode RMSSTD mengukur kehomogenan dari kelompok yang terbentuk pada setiap tahap. Semakin kecil nilai RMSSTD semakin tinggi nilai kehomogenan kelompok yang terbentuk pada tahap tersebut. Masing-masing data kredit, tunai, laba akan diuji

menggunakan metode RMSSTD untuk mencari cluster yang paling optimum.



Gambar 12 Grafik hasil metode RMSSTD

Gambar 11 terlihat bahwa pada pengujian RMSSTD pada ketiga data yang dilakukan pengujian, nilai dari algoritma Fuzzy C-Means cenderung lebih baik dengan nilai pada gambar 11. Semakin kecil nilai RMSSTD semakin tinggi nilai kehomogenan kelompok yang terbentuk pada tahap tersebut.

KESIMPULAN DAN SARAN

a. Kesimpulan

Berdasarkan uraian permasalahan dan pembahasan pada bab sebelumnya, maka didapatkan beberapa kesimpulan tentang implementasi dan pengujian algoritma *Fuzzy C-Means* dan *K-Means Clustering* sebagai berikut:

1. Dalam clustering untuk pengelompokan pelanggan pada PT.Part Station Jember, data transaksi nota dilakukan preprocessing data dan mengambil atribut yang diperlukan. Data ini selanjutnya diklasterisasi menggunakan algoritma Fuzzy C-Means dan K-Means Clustering. Penentuan cluster algoritme K-Means diperoleh dari perhitungan jarak terkecil antara data dengan pusat cluster (centroid). Sedangkan penentuan cluster algoritme Fuzzy C Means diperoleh dari nilai derajat keanggotaan terbesar data terhadap setiap cluster.
2. Hasil pengujian metode Silhouette Coefficient pada data kredit menunjukkan algoritma Fuzzy C-Means lebih baik dibandingkan algoritma K-Means Clustering dengan nilai 0.61 untuk algoritma Fuzzy C-Means dan 0.59 untuk algoritma K-Means Clustering. Sedangkan pada pengujian data tunai dan laba menunjukkan algoritma K-Means Clustering lebih baik dibandingkan Fuzzy C-Means dengan nilai masing-masing 0.66 dan 0.25 untuk algoritma Fuzzy C-Means dan 0.79 dan 0.4 untuk algoritma K-Means Clustering. hal ini menunjukkan metode validasi Silhouette Coefficient ini sangat dipengaruhi oleh data yang digunakan.
3. Hasil pengujian metode RMSSTD menunjukkan algoritma Fuzzy C-Means lebih baik dibandingkan algoritma K-Means Clustering dari semua data yang dilakukan pengujian, baik itu data kredit, tunai dan laba. Dengan nilai rata-rata 0.024979667 data kredit, 0.06873552 data tunai, 0.010187823 data laba untuk algoritma Fuzzy C-Means sedangkan nilai rata-rata 0.03229311 data kredit, 0.105352232 data tunai, 0.064485279 data laba untuk algoritma *K-Means Clustering.*

b. Saran

Beberapa saran yang dapat dijadikan pertimbangan dalam mengembangkan penelitian ini adalah:

1. Pengembang selanjutnya diharapkan dapat menggunakan algoritma clustering yang berbeda seperti *Hierarchical Clustering* dan *Partitional Clustering*.
2. Dapat menggunakan metode uji validasi cluster yang berbeda seperti DBI *(Davies Building Index)* untuk mengetahui keakuratan performa.
3. Pada penelitian ini hanya menekankan performa metode clustering. Agar lebih bermanfaat pengembang selanjutnya diharapkan dapat membangun sistem aplikasi analisis performa metode clustering.

DAFTAR PUSTAKA

Adiana, B. E., Soesanti, I., & Permanasari, A. E. (2018). Analisis Segmentasi Pelanggan Menggunakan Kombinasi RFM Model dan Teknik Clustering. Jutei, 2(2), 23–32. https://doi.org/10.21460/jutei.2017.21.76

Agustini, F. (2017). Implementasi Algoritma Fuzzy C-Means Studi Kasus, 3(1), 127–132.

Algoritma, M., Ud, K. D. I., & Wijaya, D. A. (2015). CLUSTERING PEMILIHAN PELANGGAN POTENSIAL.

Atau Langsung Ke. (n.d.), 1–11.

Fitria Hastuti, N. (2013). Pemanfaatan metode k-means clustering dalam penentuan penerima beasiswa. Skripsi.

Fox, R. (2010). No {{Title}}, 5(1976), 265–288.

Ramadhan, F. R. (2017). Analisis Perbandingan Metode Clustering Single Linkage Dan K-Means Pada Pengelompokan Tugas Akhir Teknik Informatika. Laporan Tugas Akhir.

Syarif, R., Furqon, M. T., & Adinugroho, S. (2018). Perbandingan Algoritme K-Means Dengan Algoritme Fuzzy C Means (FCM) Dalam Clustering Moda Transportasi Berbasis GPS, 2(10), 4107–4115.

Tanzil Furqon, M., & Muflikhah, L. (2016). Clustering the Potential Risk of Tsunami Using Density-Based Spatial Clustering of Application With Noise (Dbscan). Journal of Enviromental Engineering and Sustainable Technology, 3(1), 1–8. https://doi.org/10.21776/ub.jeest.2016.003.01.1

Yohanes. (2016). Analisis Perbandingan Algoritma Fuzzy C-Means dan K-Means. Annual Research Seminar 2016, 2(1), 151–155.

Yudiantri. (2017). Pengenalan Software R. Time Series Forecasting with R: From

Classical to Modern Methods, 1–9. Retrieved from http://matematika.fmipa.unand.ac.id/images/bahan-seminar/pakyudi.pdf