

## CLUSTERING LOYALITAS PELANGGAN DENGAN METODE RFM (RECENCY, FREQUENCY, MONETARY) DAN FUZZY C-MEANS

**Sudriyanto**

Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Nurul Jadid  
JL. KH Zaini Mun'im, Pondok Pesantren Nurul Jadid, PO.Box. 1 Paiton Probolinggo 67291  
\*Email: sudriyanto21@gmail.com

### Abstrak

*Semakin maju dan berkembangnya dunia usaha, menciptakan persaingan yang luar biasa ketata. Persaingan yang ada mengharuskan pemilik usaha untuk selalu dapat memahami sebuah pasar yang terus berkembang. Banyak faktor yang mempengaruhi berkembangnya usaha, salah satunya pelanggan. Untuk membangun strategi yang handal dalam dunia usaha, harus menemukan cara untuk menarik dan mengembangkan aset manusia yang tepat dan menjaga mereka. Untuk mengetahui hal tersebut, dilakukan penelitian mencari pola dari mutu calaster dengan pemilihan atribut menggunakan metode RFM (Recency, Frequency, Monetary) untuk mengklaster pelanggan yang lebih efisien dan akurat dengan algoritma Fuzzy C-Means untuk memilih pelanggan yang potensial dan loyal. Dari hasil pengukuran validitas menggunakan Partition Coefficient Index (PCI) dan Xie dan Beni (XBI) 5 cluster dengan pangkat dua dan tiga dengan nilai PCI pangkat dua 0,8156, pangkat tiga 0,5860, untuk nilai XBI pangkat dua 0,0069, pangkat tiga 0,0632 dimana nilai PCI yang mendekati angka satu mempunyai arti kualitas cluster yang didapat semakin baik, sedangkan untuk XBI yang semakin kecil mempunyai arti kualitas hasil pengelompokan yang semakin baik. Dari hasil pengukuran validitas menggunakan PCI dan XBI, nilai akurasi FCM dengan pangkat dua lebih baik dibandingkan dengan nilai akurasi FCM dengan pangkat tiga.*

**Kata kunci:** RFM (Recency, Frequency, Monetary), Fuzzy C-Means, clustering, validitas cluster, Partition Coefficient Index (PCI) dan Xie dan Beni (XBI)

### 1. PENDAHULUAN

Semakin maju dan berkembangnya dunia usaha, menciptakan persaingan yang luar biasa ketat. Persaingan yang ada mengharuskan pemilik usaha untuk selalu dapat memahami pasar yang terus berkembang. Dalam pengelolaan usaha, ada beberapa faktor yang membuat usaha itu maju dan berkembang pesat. Banyak faktor yang mempengaruhi berkembangnya usaha, salah satunya adalah pelanggan. Pelanggan merupakan faktor yang menentukan maju dan tidaknya suatu usaha bisnis. Untuk membangun strategi handal dalam dunia usaha, harus menemukan cara untuk menarik dan mengembangkan aset manusia yang tepat dan menjaga mereka.

Pertumbuhan internet secara cepat dan ini berkaitan dengan peluang untuk memasarkan teknologi yang sangat meningkat dan perubahan cara pandang hubungan antara konsumen dengan perusahaan dan para pengusaha mengatur konsumen. Kepuasan pelanggan dan daya loyalitas pelanggan merupakan hal yang sangat penting dalam perkembangan usaha. Kepuasan pelanggan adalah tingkat dimana anggapat terhadap produk sesuai dengan harapan seseorang. Ketika pelanggan fokus terhadap organisasi bisnis baik kedalam organisasi maupun keluar organisasi.

Perkembangan internet yang begitu cepat memberikan limpahan data yang sangat besar, meskipun belum didapatkan manfaat data yang besar ini. Praktisi database internet dan peneliti baru-baru ini menyatakan minat praktisi bisnis yang besar pada database pemasaran.

Selama ini pendekatan pemasaran yang ada adalah berorientasi kepada produk, hal ini menyebabkan banyak pelaku bisnis hanya fokus pada produk yang mereka buat. Pendekatan produk ternyata berdampak sulitnya mengidentifikasi pelanggan secara benar.

Perusahaan ini perihatin dengan meningkatnya nilai pelanggan melalui analisis siklus hidup pelanggan. Hal ini merubah cara pandang perusahaan terhadap siklus perusahaan. Jika dahulu perusahaan memiliki cara pandang berorientasi produk, kini diganti oleh cara pandang berorientasi pelanggan. Pelanggan adalah masalah utama bagi semua organisasi yang ingin menciptakan dan menjaga keunggulan kompetitif di dunia yang sangat kompetitif ini.

Pemanfaatan data mining untuk menemukan pola-pola yang terpendam dari data yang besar mulai digunakan. Klastering merupakan salah satu peran utama dalam data mining. Terdapat

banyak teknik klustering data mining seperti yang tercatat. Tujuan dari clustering memprediksi secara akurat memprediksi kelas target untuk setiap record data baru, yaitu data yang tidak dalam data historis.

Data konsumen dan perkembangan teknologi adalah sumber untuk mengolah hubungan antar konsumen dengan perusahaan [2]. Pengolahan CRM tersebut meliputi klasifikasi dan klustering. Klasifikasi bertujuan untuk memetakan pelanggan berdasarkan geografis dan kedekatan dengan sumber industri [3]. Sedangkan penggunaan clustering pada pengolahan CRM untuk mengklusterkan pelanggan berdasarkan interaksi yang terjadi antara pelanggan dengan perusahaan.

Data pelanggan cenderung merupakan data transaksi pelanggan yang memiliki banyak atribut. Adanya atribut ini berakibat pada pengolahan data yang kurang baik [6]. Perlu adanya pemilihan atribut yang baik untuk mendapatkan hasil yang lebih optimal. Untuk mengatasi beberapa kekurangan dalam pemilihan atribut yang ada, diusulkan penggunaan metode RFM (*Recency, Frequency dan Monetary*). *Recency* adalah jarak terlama dari data itu diambil. Ini mengandung pengertian bahwa jarak konsumen terakhir kalinya melakukan transaksi kepada organisasi dihitung dari data diambil. Satuan yang biasa digunakan adalah perhari, perbulan atau pertahun. Semakin dekat dengan pengambilan data, maka pelanggan tersebut semakin memiliki point tinggi. *Frequency* adalah seberapa sering pelanggan melakukan transaksi kepada organisasi. Semakin sering melakukan transaksi ini memungkinkan bahwa pelanggan tersebut merupakan pelanggan potensial. Transaksi yang dimaksud tidak tergantung pada berapa banyak yang pelanggan beli, tetapi satu transaksi pun bisa mewakili bahwa pelanggan tersebut potensial. *Monetary* adalah seberapa banyak uang yang pelanggan habiskan untuk melakukan transaksi kepada organisasi, semakin tinggi nilai mata uang yang dihabiskan maka, semakin tinggi pula potensial pelanggan tersebut

Untuk mengatasi masalah pemilihan atribut pelanggan diusulkan menggunakan model RFM, sedangkan untuk proses klustering menggunakan algoritma Fuzzy C-Means. Pemilihan metode RFM adalah karena metode yang efektif untuk segmentasi pelanggan dan objektifitasnya baik [10]. Metode RFM bertumpu pada pemilihan tiga atribut. Fuzzy C-Means merupakan salah satu metode algoritma yang terkenal untuk analisis kluster dan telah digunakan secara luas di berbagai bidang termasuk data mining, analisis data statistik dan aplikasi bisnis lainnya.

Analisis *Cluster* adalah teknik statistik yang digunakan untuk mengidentifikasi satu set kelompok yang kedua meminimalkan dalam variasi kelompok dan memaksimalkan antara kelompok variasi berdasarkan fungsi jarak atau perbedaan, dan tujuannya adalah untuk menemukan set optimal *cluster*. Berdasarkan atribut RFM dan metode Fuzzy C-Means untuk mengelompokkan nilai pelanggan.

## 2. METODOLOGI

Dalam penelitian ini akan dilakukan beberapa tahapan eksperimen menggunakan tahapan yang dilakukan dalam sebuah proses penelitian berikut :

### 1.1. Pengumpulan Data

Data uji yang dilakukan pada penelitian ini bersumber pada sebuah database penjualan UD. Budi Luhur Kota Proboinggo. Tahap ini bertujuan untuk mengumpulkan berbagai informasi mengenai penelitian dengan cara wawancara dan dokumentasi

Data yang digunakan adalah sebagai berikut :

1. Data pelanggan yang melakukan transaksi dengan UD. Budi Luhur Kota Proboinggo baik dalam jumlah kecil maupun besar dalam masa waktu/periode yang ditentukan
2. Data historikal penjualan produk dengan atribut sesuai dengan yang dibutuhkan oleh model RFM, terdiri dari:
  - a. Waktu transaksi terakhir
  - b. Jumlah nominal transaksi
  - c. Jumlah frekuensi transaksi
3. Data historikal penjualan produk yang digunakan adalah data transaksi untuk segala macam produk seperti, unilever, wings, cussen, indofood dan lain - lain.

### 1.2. Praproses Data

Data mentah yang diperoleh dari data pelanggan dan data historikal penjualan pada studi kasus akan melalui praproses data yang meliputi :

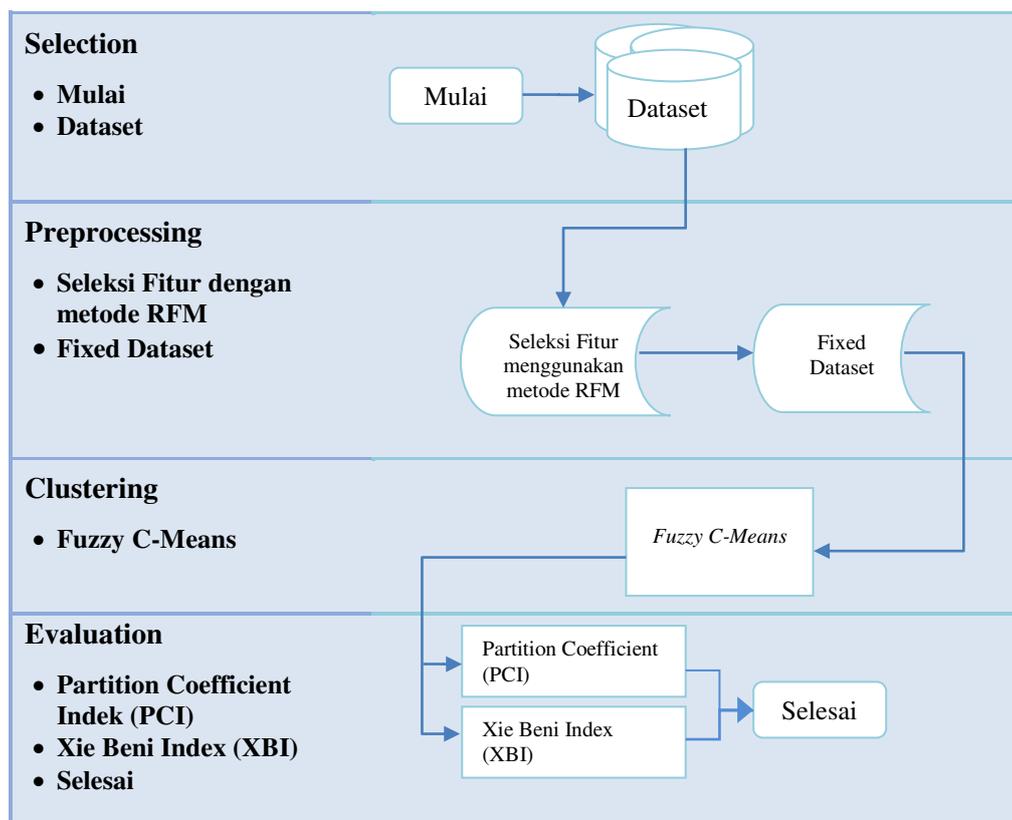
1. Proses pembersihan data untuk menghilangkan baris data dengan nilai kosong atau data yang tidak valid (*data cleaning and integration*)
2. Pemilihan data berdasarkan atribut yang disesuaikan dengan model RFM yaitu: atribut waktu transaksi, jumlah frekuensi transaksi dan jumlah nominal transaksi untuk setiap pelanggan (*data selection*)
3. Persiapan data dengan cara mereduksi kolom data yang tidak sesuai dengan kebutuhan dalam melakukan proses klasterisasi (*data preprocessing*)
4. Transformasi data kedalam bentuk yang terukur sehingga dapat digunakan sebagai atribut untuk proses klasterisasi (*transformation*)

### 1.3. Perhitungan Nilai RFM

Berpedoman pada model RFM yang terdiri dari tiga kriteria yaitu *recency*, *frequency*, *monetary*, maka dilakukan ekstraksi data dengan atribut-atribut yang dibutuhkan untuk mengukur nilai loyalitas pelanggan. Data yang digunakan adalah data yang telah melewati tahap pengumpulan dan praproses data. Selanjutnya dilakukan proses klasterisasi.

1. *Recency* → Tanggal acuan terakhir transaksi (31/12/2016) periode 2016
2. *Frequency* → Jumlah transaksi selama periode 2016
3. *Monetary* → Nilai transaksi (uang) selama periode 2016

### 1.4. Metode yang diusulkan



Gambar 2.1. Alur Tahapan Metode yang Diusulkan

### 3. HASIL DAN PEMBAHASAN

#### 1.5. Pengumpulan data

Data uji yang dikumpulkan pada penelitian ini bersumber pada sebuah database penjualan pada UD. Budi Luhur yang berada di kota Probolinggo. Data yang diambil dari gabungan beberapa tabel seperti data pelanggan, data transaksi penjualan. Periode data yang diambil dari data transaksi penjualan selama tahun 2016.

#### 1.6. Praproses data

Tahapan implementasi ini dimulai dari proses pengumpulan data yang telah dijelaskan di atas. Data yang digunakan adalah data pelanggan dan data historikal transaksi penjualan untuk periode 1 Januari 2016 – 31 Desember 2016. Sebelum dilakukan praproses data, data mentah yang dikumpulkan dari pusat grosir adalah data-data dengan atribut ditunjukkan dalam tabel berikut :

**Tabel 3.1 Tabel Pelanggan**

No	Nama Atribut	Keterangan
1	KdPelanggan	Kode Pelanggan Yang pernah melakukan transaksi
2	Nama Pelanggan	Nama pelanggan yang pernah melakukan transaksi dengan perusahaan

**Tabel 3.2 Tabel Transaksi Penjualan**

No	Nama Atribut	Keterangan
1	TgJual	Tanggal pada saat pelanggan melakukan transaksi pembelian (periode analisis 1 Januari 2016 s/d 31 Desember 2016)
2	NoJual	Nomer urut transaksi pelanggan
3	KdPelanggan	Kode pelanggan yang melakukan transaksi
4	Nama Pelanggan	Nama pelanggan yang melakukan transaksi
5	KdBarang	Kode barang yang dibeli pelanggan berdasarkan jumlah
6	Nama Barang	Nama barang yang dibeli pelanggan berdasarkan jumlah
7	Jumlah	Jumlah satuan barang yang dibeli pelanggan
8	HrgSatuan	Harga satuan per barang yang dibeli pelanggan
9	Nilai	jumlah harga produk yang dibeli pelanggan

Setelah itu dilakukan praproses data kepada tabel transaksi dengan mereduksi atribut yang tidak dibutuhkan dalam model RFM. Sehingga mendapat hasil tabel sebagai berikut :

**Tabel 3.3 Hasil pemilihan data dari tabel transaksi**

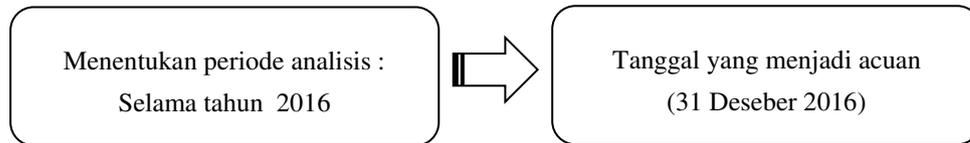
No	Nama Atribut	Keterangan
1	TgJual	Tanggal pada saat pelanggan melakukan transaksi pembelian (periode analisis 1 Januari 2015 s/d 1 Desember 2015)
2	NoJual	Nomer urut transaksi pelanggan
3	KdPelanggan	Kode pelanggan yang melakukan transaksi
4	Nama Pelanggan	Nama pelanggan yang melakukan transaksi
5	Nilai	jumlah harga produk yang dibeli pelanggan

### 1.7. Penilaian Pelanggan Berbasis Model RFM

Pada tahap ini, dilakukan pemilihan atribut yang disesuaikan dengan kebutuhan kriteria model RFM, yaitu rentang waktu transaksi akhir pelanggan dengan periode analisis, jumlah frekuensi transaksi yang dilakukan pelanggan selama periode analisis, serta jumlah nominal transaksi untuk setiap pelanggan selama periode analisis. Berikut uraian mengenai atribut yang dibutuhkan untuk model RFM :

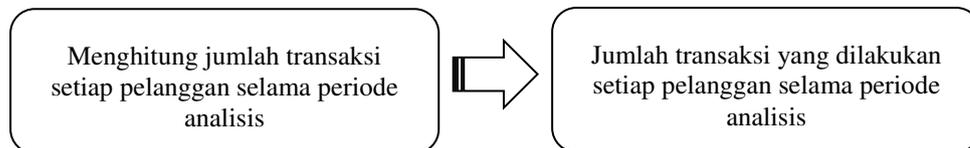
1. Kriteria R (recency)

Mebutuhkan atribut yang menunjukkan adanya rentang waktu transaksi terakhir pelanggan dengan periode analisis, sehingga atribut yang dibutuhkan adalah atribut tanggal transaksi



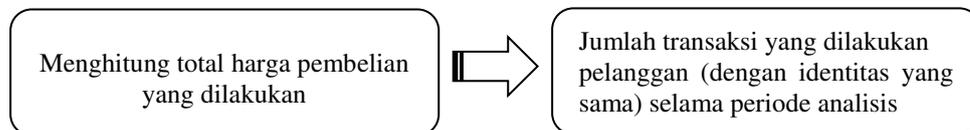
2. Kriteria F (frequency)

Mebutuhkan atribut yang merepresentasikan berapa kali pelanggan melakukan transaksi. Kriteria ini dapat dilihat dari berapa banyak pelanggan dengan nama yang sama muncul dalam data transaksi. Atribut yang dibutuhkan adalah atribut tanggal transaksi yang dihitung jumlahnya.



3. Kriteria M (monetary)

Mebutuhkan atribut yang berhubungan dengan harga yang telah dihabiskan pelanggan selama melakukan transaksi, sehingga atribut yang dibutuhkan adalah atribut harga total yang harus dibayar pelanggan.



Berikut adalah data yang telah melewati tahap pengumpulan dan praproses data dan selanjutnya dilakukan proses klusterisasi:

**Tabel 3.4 Cuplikan Hasil Transformasi Data ke dalam Recency, Frequency, Monetary**

No	Kode Pelanggan	Nama Pelanggan	Recency	Frequency	Monetary
01	GD006	HASAN	29	31	18.552.549
02	GD011	PAK RONI	3	93	10.254.430
03	GD014	TOKO FERI	10	25	17.635.540
04	GD020	FREDI (REFISI)	14	13	19.778.155
05	GD021	KARNO	59	13	7.127.352
06	GD022	BAHAK UDIN	6	22	5.939.172
07	GD023	YUDI (SALES)	29	16	5.971.286
08	GD028	YUDI SOPIR	4	39	16.816.855
09	GD042	UD SEHATI JAYA	10	40	69.552.000
10	GND001	TOKO SONY JAYA	16	20	10.208.378
....	....	....	....	....	....
502	WNS09	TOKOK RAMA	5	9	12.161.251
503	WNS10	TOKOK AGEN SOSIS	14	14	7.341.213

**1.8. Proses Clustering**

Pada tahap ini dilakukan proses klasterisasi pelanggan menggunakan metode *Fuzzy C-Means*. Data yang digunakan dalam proses ini adalah keluaran tahap analisa model RFM. Proses clustering menggunakan algoritma FCM di ujicoba menggunakan 5 cluster dengan pangkat 2 dan 3. Berikut hasil ujicoba dengan pangkat dua dan tiga :

1. Jumlah cluster = 5
2. Pangkat = 2
3. Maksimum iterasi = 100
4. Error terkecil yang diharapkan =  $10^{-5}$
5. Fungsiobjektif awal = 0
6. Iterasi awal = 1

**Tabel 3.5 Hasil akhir derajat keanggotaan pada tiap cluster dengan pangkat 2 (dua)**

Data Ke	Derajat keanggotaan ( $\mu$ ) data pada Cluster ke-					MAX	Data Cenderung Masuk ke Cluster ke-				
	1	2	3	4	5		c1	c2	c3	c4	c5
1	0,979	0,000	0,000	0,018	0,002	0,979	x				
2	0,683	0,004	0,001	0,292	0,020	0,683	x				
3	0,051	0,236	0,008	0,095	0,610	0,610					x
4	0,052	0,411	0,010	0,091	0,437	0,437					x
5	0,086	0,023	0,002	0,381	0,508	0,508					x
6	0,592	0,004	0,001	0,380	0,023	0,592	x				
7	0,981	0,001	0,000	0,015	0,003	0,981	x				
8	0,997	0,000	0,000	0,003	0,000	0,997	x				
9	0,015	0,009	0,001	0,046	0,930	0,930					x
10	0,989	0,000	0,000	0,009	0,001	0,989	x				
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
501	0,955	0,002	0,000	0,036	0,007	0,955	x				
502	0,982	0,000	0,000	0,016	0,002	0,982	x				
503	0,984	0,001	0,000	0,014	0,002	0,984	x				
<b>Jumlah :</b>							<b>288</b>	<b>22</b>	<b>4</b>	<b>139</b>	<b>50</b>

**Tabel 3.6 Hasil akhir derajat keanggotaan pada tiap cluster dengan pangkat 3 (tiga)**

Data Ke	Derajat keanggotaan ( $\mu$ ) data pada Cluster ke-					MAX	Data Cenderung Masuk ke Cluster ke-				
	1	2	3	4	5		c1	c2	c3	c4	c5
1	0,713	0,195	0,009	0,026	0,058	0,713	x				
2	0,398	0,459	0,013	0,038	0,091	0,459		x			
3	0,124	0,160	0,046	0,033	0,341	0,341					x
4	0,114	0,144	0,047	0,417	0,278	0,417				x	
5	0,131	0,233	0,021	0,074	0,537	0,537					x
6	0,350	0,509	0,012	0,037	0,092	0,509		x			
7	0,880	0,074	0,004	0,013	0,027	0,880	x				
8	0,813	0,124	0,007	0,018	0,039	0,813	x				
9	0,015	0,009	0,001	0,046	0,930	0,930					x
10	0,989	0,000	0,000	0,009	0,001	0,989	x				
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
502	0,723	0,187	0,009	0,025	0,056	0,723	x				
503	0,893	0,066	0,004	0,012	0,025	0,893	x				
<b>Jumlah :</b>							<b>255</b>	<b>157</b>	<b>4</b>	<b>24</b>	<b>63</b>

### 1.9. Evaluasi PCI dan XBI

Untuk mengetahui validasi hasil eksperimen digunakan uji validasi *partition coefficient index* (PCI) dan Xie dan Beni (XBI) dimana nilai PCI dalam rentang (0,1), nilai yang semakin besar mendekati 1 (satu) mempunyai arti kualitas cluster yang didapat semakin baik, sedangkan nilai XBI dimana nilai indeks yang semakin kecil mempunyai arti kualitas hasil pengelompokan yang semakin baik (Wu dan Yang 2005). Berikut hasil pengelompokan yang didapat :

**Tabel 3.7 Validasi hasil klasterisasi dengan PCI dan XBI**

FCM dengan Pangkat	Partition Coefficient Index (PCI)	Xie Beni Index (XBI)
2	0,8156	0,0069
3	0,5860	0,0632

## 4. KESIMPULAN

Dari penelitian yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

Dengan menggunakan metode RFM *Recency*, *Frequency* dan *Monetary* untuk pemilihan atribut dan Algoritma Fuzzy C-Means untuk proses pengklasteran lebih objektif dalam meneliti loyalitas pelanggan serta mempermudah pemilik usaha sebagai dasar pengambilan keputusan dalam pemberian reward kepada pelanggan agar lebih tepat sasaran, dari hasil pengukuran validitas menggunakan *Partition Coefficient Index* (PCI) dengan FCM pangkat dua 0,8156 dan pangkat tiga 0,5860 dimana nilai yang semakin besar (mendekati 1) mempunyai arti bahwa kualitas cluster yang didapat semakin baik. Sedangkan pengukuran Xie Beni Index (XBI) dengan FCM pangkat dua 0,0069 dan pangkat tiga 0,0632 dimana nilai indeks yang semakin kecil mempunyai arti kualitas hasil pengelompokan yang semakin baik. Dari hasil pengukuran validasi *Coefficient Index* (PCI) dan Xie Beni Index (XBI), nilai akurasi *Fuzzy C-Means* dengan pangkat 2 (dua) lebih baik dibandingkan dengan nilai akurasi FCM dengan pangkat 3 (tiga).

## DAFTAR PUSTAKA

- T. Kristanto and . A. and Rachman, "Analisa Data Mining Metode Fuzzy untuk Customer Relationship Management pada Perusahaan Tour & Travel.," *SESINDO* , 2013.
- E. Ngai, L. Xiu and D. Chau, "Application of data mining techniques in customer relationship management:," *Expert Systems with Applications*, p. 2592–2602, 2009.
- C. Rygielski, J.-C. Wang and D. C. Yen, "Data mining techniques for customer," *Technology in Society* 24, p. 483–502, 2002.
- H. Pahlevie and T. Safiria Ramadhani, "Module 5: Customer Relationship Management," in *Institut Teknologi Sepuluh Nopember*, surabaya, 2012.
- J. R. Fonseca, "Customer satisfaction study via a latent segment model.," *Journal of Retailing and Consumer Services* 16.5 , pp. 352-359., 2009.
- Y.-H. Hu and T.-W. Yeh, "Discovering valuable frequent patterns based on RFM analysis without," *Knowledge-Based Systems*, p. 76–88, 2014.
- J. Wu and Z. Lin, "Research on Customer Segmentation Model," p. 316–318, 2005.
- K.-. L. Wu, "Analysis of parameter selections for fuzzy c-means," *Pattern Recognition* 45, p. 407–415, 2012.
- Y. Tu and Z. Yang, "An enhanced Customer Relationship Management classification framework," *Expert Syst. Appl*, p. 2137–2146, 2013.
- A. Aktepe, S. Ersöz and B. Toklu, "Customer satisfaction and loyalty analysis with classification algorithms," *Computers & Industrial Engineering* 86, p. 95–106, 2015.
- D. Zheng, "Application of Silence Customer Segmentation in," *Journal of Information & Computational Science*, p. 4337–4347, 2013.
- A. Sheshasayee and P. Sharmila, "Comparative Study of Fuzzy C Means and K Means," *Indian Journal of Science and Technology*, Vol 7(6), p. 853–857, 2014.
- C.-H. Cheng and Y.-S. Chen, "Classifying the segmentation of customer value via RFM model and RS theory," *Expert Systems with Applications*, p. 4176–4184, 2009.

- J. A. McCarty and M. Hastak, "Segmentation approaches in data-mining: A comparison of," *Journal of Business Research*, p. 656 – 662, 2007.
- F. A. Hermawati, "Data Mining," in *CV ANDI OFFSET*, Yogyakarta, 2013.
- I. I. John Wiley & Sons, "Manajemen Sumber Daya Manusia Untuk Perusahaan," in *PT. Raja Grafindo Persada*, Jakarta, 2002.
- H. Simamora, "Manajemen Sumber Daya Manusia," in *STIE YKPN*, Yogyakarta, 2004, p. 514.
- Mahmudi, "Manajemen Kinerja Sektor Publi," in *Unit Penerbit dan Percetakan Akademi Manajemen Perusahaan YKPN*, Yogyakarta, 2005, p. 89.
- M. CHANDRASHEKARAN, K. ROTTE, S. S. TAX and R. GREWAL, "Satisfaction Strength and Customer Loyalty," *Journal of Marketing Research*, p. 153–163, 2007.
- S. K. a. H. Pornomo, "Aplikasi Logika Fuzzy untuk Pendukung Keputusan," in *GRAHA ILMU*, Yogyakarta, 2010.