

## OPTIMIZATION OF MARKET BASKET ANALYSIS USING CENTROID-BASED CLUSTERING ALGORITHM AND FP-GROWTH ALGORITHM

Fitri Nuraeni<sup>\*1</sup>, Dewi Tresnawati<sup>2</sup>, Yoga Handoko Agustin<sup>3</sup>, Gisna Fauzian Dermawan<sup>4</sup>

<sup>1,2,3,4</sup> Teknik Informatika, Jurusan Ilmu Komputer, Fakultas Teknik, Institut Teknologi Garut  
Email: <sup>1</sup>[fitri.nuraeni@itg.ac.id](mailto:fitri.nuraeni@itg.ac.id), <sup>2</sup>[dewi.tresnawati@itg.ac.id](mailto:dewi.tresnawati@itg.ac.id), <sup>3</sup>[yoga.handoko@itg.ac.id](mailto:yoga.handoko@itg.ac.id), <sup>4</sup>[1806122@itg.ac.id](mailto:1806122@itg.ac.id)

(Naskah masuk: 21 Juni 2022, Revisi: 12 Juli 2022, diterbitkan: 26 Desember 2022)

### Abstract

The proliferation of the food and beverage sales business requires the creativity of business owners to offer their flagship products to every consumer, both new and subscribed consumers. A large number of menu choices makes the ordering process long because consumers are confused about which menu will be the best choice. the seller to be able to provide the right recommendations so that orders can take place faster. Shopping cart analysis is an activity that has often been done to find out the items found that are sold simultaneously. The FP-Growth association method is a faster algorithm for generating association rules, but the association process in large dataset sizes tends to add large items so that the accuracy value of association rules decreases. So that in this study, the grouping of datasets was carried out using a clustering model with a centroid-based algorithm, namely k-means, k-medoids, and fuzzy c-means. This research was conducted through dataset collection, dataset preparation, clustering modeling, evaluation of clustering models using DBI and silhouette index, association modeling, and evaluation of association models using lift ratio. The results of this study showed that the clustering model with the best DBI and silhouette index values was at k=3 for k-means, k=2 for k-medoids, and k=7 for fuzzy c-means. The number of association rules is generated from the grouped data set using fuzzy c-means, but the highest average lift ratio is in the association rules generated from the grouping data set using k-means. From the association model using k-means and FP-Growth, 32 unique association rules were found with the 4 most frequently found items, namely cireng chili oil, regal milk coffee, banana cheese, and vietnam drip.

**Keywords:** centroid-based algorithm, clustering, FP-Growth, lift ratio, market basket analysis.

## OPTIMALISASI ANALISIS KERANJANG PASAR MENGGUNAKAN ALGORITMA CENTROID-BASED CLUSTERING DAN ALGORITMA FP-GROWTH

### Abstrak

Makin menjamurnya bisnis dibidang penjualan makanan dan minuman, menuntut kreatifitas pemilik bisnis untuk menawarkan produk andalannya pada setiap konsumen, baik konsumen baru maupun yang sudah berlangganan. Banyaknya pilihan menu membuat lama proses order, karena konsumen kebingungan memilih menu mana yang akan menjadi pilihan terbaik. Penjual dituntut untuk bisa memberikan rekomendasi yang tepat agar proses order dapat berlangsung lebih cepat. Analisis keranjang belanja merupakan kegiatan yang sudah sering dilakukan untuk mengetahui pasangan item barang yang sering terjual bersamaan. Metode asosiasi FP-Growth merupakan algoritma yang lebih cepat dalam menghasilkan aturan asosiasi, namun proses asosiasi pada ukuran dataset yang besar cenderung mengabaikan *itemset* dalam ukuran besar sehingga nilai akurasi aturan asosiasi berkurang. Sehingga pada penelitian ini, dilakukan pengelompokan dataset yang memiliki kemiripan menggunakan model *clustering* dengan algoritma *centroid-based* yaitu *k-means*, *k-medoids* dan *fuzzy c-means*. Penelitian ini dilakukan melalui tahapan pengumpulan dataset, persiapan dataset, pemodelan *clustering*, evaluasi model *clustering* menggunakan DBI dan *silhouette index*, dilanjutkan pemodelan asosiasi dan evaluasi model asosiasi menggunakan *lift ratio*. Hasil dari penelitian ini didapatkan bahwa model *clustering* dengan nilai DBI dan *silhouette index* terbaik berada pada k=3 untuk *kmeans*, k=2 untuk *kmedoids* dan k=7 untuk *fuzzy c-means*. Jumlah aturan asosiasi terbanyak dihasilkan dari dataset hasil *clustering* menggunakan *fuzzy c-means*, namun rata-rata *lift ratio* tertinggi berada pada aturan asosiasi yang dihasilkan dari dataset hasil *clustering* menggunakan *k-means*. Dari model asosiasi menggunakan *k-means* dan *FP-Growth* didapatkan 32 aturan asosiasi yang unik dengan 4 item yang paling sering muncu yaitu cireng chili oil, kopi susu regal, pisang keju dan vietnam drip.

**Kata kunci:** algoritma centroid-based, analisis keranjang pasar, clustering, FP-Growth, lift ratio.

## 1. PENDAHULUAN

Walau dihantam pandemi covid-19 mulai tahun 2020 sampai saat ini, usaha di bidang penjualan makanan dan minuman merupakan bidang usaha yang dapat bertahan dan tidak pernah mati[1]. Proses penjualan makanan dan minuman banyak beralih menggunakan sistem penjualan *online*[2] disamping tetap mempertahankan gerai-gerai penjualan langsung. Makanan dan minuman merupakan kebutuhan pokok masyarakat, sehingga usaha bidang penjualan makanan dan minuman menjadi bisnis yang menjanjikan keuntungan serta menjadi solusi bagi yang diPHK untuk mencari penghasilan[3].

Makin menjamurnya usaha bidang penjualan makanan dan minuman, menuntut kreatifitas pelaku usaha dalam menawarkan produk andalannya pada setiap konsumen, baik konsumen baru maupun yang sudah berlangganan. Banyaknya pilihan menu membuat lama proses order, karena konsumen kebingungan memilih menu mana yang akan menjadi pilihan terbaik. Penjual dituntut untuk bisa memberikan rekomendasi yang tepat agar proses order dapat berlangsung lebih cepat, namun konsumen puas dengan menu tersebut. Selain itu, menu yang paling diminati konsumen akan menyebabkan permintaan persediaan bahan meningkat. Sehingga, penjual harus mengetahui menu apa yang paling diminati agar dapat menjaga stok bahannya.

Penjual yang sudah memiliki banyak riwayat transaksi penjualan saat ini, masih menyimpan data-data tersebut untuk laporan keuangan saja. Data yang banyak itu perlu digali untuk menghasilkan sebuah pengetahuan yang mungkin berguna dalam menjalankan usahanya[4].

Analisis keranjang pasar merupakan salah satu kegiatan data mining yaitu aturan asosiasi (*association rule*)[5]. Analisis keranjang pasar merupakan kegiatan yang sudah sering dilakukan untuk mengetahui pasangan item barang yang sering terjual bersamaan[6]. Kegiatan ini menghasilkan informasi yang dapat digunakan untuk mengetahui item barang yang paling diminati[7], mengatur tata letak item barang yang dijual[8], pengelolaan stok barang[9], sampai pada pemberian diskon untuk pembelian pasangan item[6] dan bahan promosi untuk meningkatkan profit serta strategi penjualan lainnya[10].

Untuk melakukan analisis keranjang belanja dapat digunakan algoritma popular seperti algoritma *apriori* dan *Frequent Pattern-Growth* (FP-Growth). Algoritma *apriori* merupakan algoritma yang dapat menemukan item set berdasarkan nilai minimum *support*, sehingga dapat diketahui pola stok barang berdasarkan kebutuhan permintaan pelanggan[11]. Namun, untuk efisiensi waktu penggalian data dan penyimpanan di database, maka muncul FP-Growth[12]. Algoritma FP-Growth ini dapat menentukan *frequent item set* (data yang paling

sering muncul) tanpa menggunakan generasi kandidat[13].

Permasalahan dalam analisis keranjang pasar ini tidak sekedar menentukan algoritma asosiasi yang cocok, tetapi sering kali item yang direkomendasikan kurang tepat dan *item-set* yang besar cenderung diabaikan, sehingga untuk dataset yang besar hasil rekomendasi kurang akurat[14]. Karena data yang besar dan bervariasi, maka sebelum melakukan analisis keranjang belanja dilakukan proses *clustering* terlebih dahulu agar dapat menghasilkan aturan asosiasi yang lebih banyak[15].

Pada tahun 2017, sebuah penelitian melakukan *clustering* data transaksi yang besar pernah diujicobakan pada *Self Organizing Map* (SOM) dan *K-Medoids*, menghasilkan 5 *cluster* yang kemudian menggunakan FP-Growth dan menghasilkan aturan asosiasi yang valid berdasarkan *lift ratio*[12].

Tahun 2019, data transaksi dari sebuah minimarket diproses menggunakan algoritma *k-means* dan FP-Growth, dihasilkan aturan asosiasi terbanyak pada pembagian data ke dalam 10 kelompok[15]. Pada tahun yang sama, data transaksi dari perusahaan retail diproses menggunakan *K-Medoids* menghasilkan 3 cluster dengan nilai DBI yang baik dan didapat aturan asosiasi yang dapat rekomendasikan dengan menerapkan FP-Growth[14].

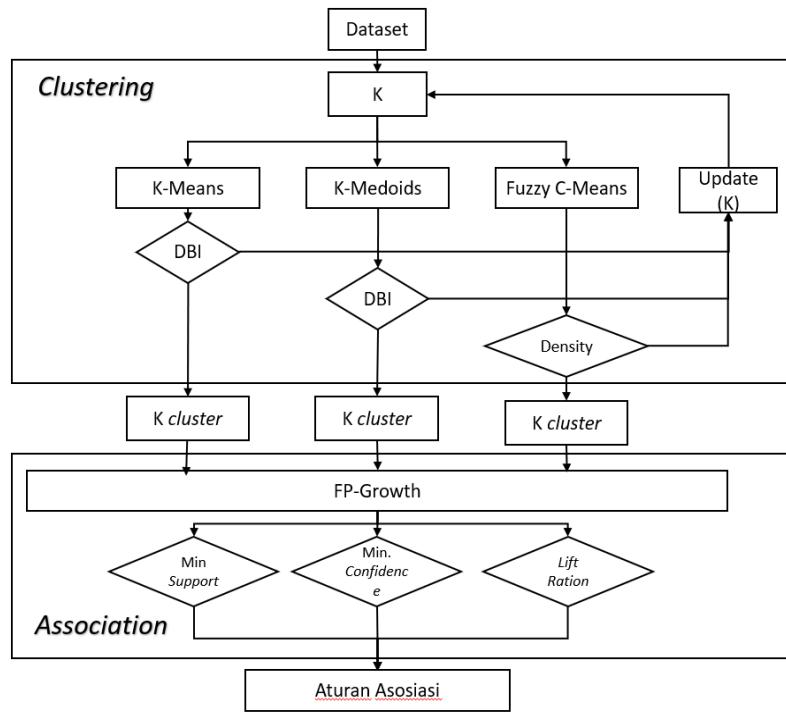
Penelitian yang serupa di tahun 2020, menggunakan data transaksi penjualan roti dan kombinasi algoritma *K-means* dan FP-Growth, menghasilkan 3 *cluster* dan pembangkitan aturan asosiasi menghasilkan *confidence* dan *lift value* yang baik[16]. Sedangkan pada tahun 2021, data transaksi perusahaan furnitur menggunakan *K-means* menghasilkan pembagian data kedalam 5 *cluster*, kemudian menggunakan FP-Growth sehingga dihasilkan aturan asosiasi yang dapat direkomendasikan[17].

Berdasarkan penelitian-penelitian sebelumnya, maka penelitian ini menggunakan algoritma FP-Growth sebagai pilihan algoritma asosiasi yang dapat menghasilkan aturan asosiasi yang efisien dalam proses dan penyimpanan datanya. Sedangkan untuk proses *clustering* akan digunakan algoritma yang setipe yaitu *centroid clustering* yaitu *k-means*, *k-medoids*, dan ditambah oleh *fuzzy c-means*. Tujuan dari penelitian ini adalah untuk mencari jumlah *cluster* yang ideal pada proses *clustering* masing-masing algoritma, kemudian mencari aturan asosiasi dari tiap *cluster* dengan nilai *lift ratio* yang baik. Selain dari itu, untuk mengetahui *cluster* hasil pemodelan algoritma *centroid clustering* mana yang dapat menghasilkan jumlah aturan asosiasi dengan *minimum support* dan *confidence*, yang terbanyak. Hasil dari penelitian ini diharapkan dapat memberikan rekomendasi bagi pemilik data untuk mengelola perusahaannya khusus terkait penawaran menu dan diskon untuk pembelian pasangan item.

## 2. METODE PENELITIAN

Pada penelitian ini akan digunakan algoritma *centroid clustering* yaitu *k-means*, *k-medoids*, dan ditambah oleh *fuzzy c-means* serta algoritma asosiasi

*FP-Growth* dengan tahapan-tahapan seperti yang terlihat pada gambar 1 dibawah ini.



Gambar 1. Kerangka Penelitian

### 2.1. Tahapan Penelitian

Proses penelitian ini diawali dengan mempersiapkan dataset sesuai pada gambar 1. Dataset berasal dari data transaksi penjualan kedai kopi di daerah Kabupaten Garut yang melayani pembelian *online* maupun langsung di tempat. Dataset ini terdiri dari 4.812 record transaksi dengan 42 item barang yang terjual. Selanjutnya dilakukan data transformasi, merubah dari bentuk *record data* ke dalam bentuk *transaction data*, seperti pada gambar 2.

dicari  $k$  jumlah *cluster* yang optimal berdasarkan *davies bouldine index*[18].

*K-cluster* dari masing-masing pemodelan diproses kembali menggunakan algoritma *FP-Growth* untuk mendapatkan aturan asosiasi. Setiap aturan asosiasi akan dibandingkan berdasarkan nilai *lift ratio* untuk mengetahui validitasnya[19]. Kemudian akan dipilih aturan asosiasi yang memiliki *lift ratio*, minimal *support* dan minimal *confidence* yang tertinggi.

### 2.2. Centroid-Based Clustering Algorithm

Pada penelitian ini pemodelan pertama adalah proses *clustering*, yang bertujuan untuk mengelompokan objek-objek berdasarkan karakteristik sehingga didapatkan homogenitas internal *cluster* yang tinggi dan heterogenitas antar *cluster* yang tinggi[20]. Algoritma *Centroid-Based Clustering* (berbasis centroid) adalah algoritma *clustering* yang menggunakan titik *centroid* dalam melakukan pengelompokan objek. Dalam salah satu algoritma berbasis centroid, tema utama yang mendasari adalah aspek penghitungan ukuran jarak antara objek dari kumpulan data yang dipertimbangkan[21]. Metode yang paling dominan dari algoritma berbasis centroid diantaranya yaitu *k-means*, *k-medoids*, dan *fuzzy c-means*.

Algoritma K-means memiliki dua fase dimana pada fase pertama; *k-centroid* diidentifikasi tergantung pada nilai  $k$  yang telah dipilih secara

A	B	C	D	E	F	G	H	I	
1	trs	Tubruk,	Seasonal	Es Susu Gu	Ice Chocol	Pisang Keji	Cireng Chil	Indomie Ri	Chocolate!
2	T00003	0	0	0	0	0	0	0	0
3	T00006	0	1	0	0	0	0	0	0
4	T00007	1	0	0	0	0	0	0	0
5	T00008	0	0	0	0	0	0	0	0
6	T00009	0	0	0	0	0	0	0	0
7	T00012	0	1	0	0	0	0	0	0
8	T00015	0	0	0	0	0	0	0	0
9	T00017	0	0	0	0	0	0	0	0
10	T00021	0	0	0	0	0	0	0	0
11	T00026	0	0	0	0	0	0	0	0
12	T00031	0	0	0	0	0	0	0	0
13	T00034	0	0	0	0	0	1	0	0
14	T00035	0	0	0	0	0	0	0	0
15	T00038	0	0	0	0	0	0	0	0
16	T00043	0	0	0	0	0	0	1	0
17	T00044	0	0	0	0	1	0	0	0

Gambar 2. Tampilan dataset

Selanjutnya dilakukan proses pemodelan pertama dengan proses *clustering* menggunakan tiga (3) algoritma *centroid clustering*, yaitu *k-means*, *k-medoids* dan *fuzzy c-means*. Pada tahapan ini akan

umum ukuran jarak dihitung menggunakan jarak Euclidian. Fase kedua melibatkan penentuan centroid baru berdasarkan nilai rata-rata objek kelompok pada *cluster*. Proses perulangan dalam menemukan centroid baru sampai konvergensi terpenuhi.

Algoritma *K-Medoids* menggunakan objek representatif sebagai titik acuan dalam setiap *cluster*, untuk meminimalkan jumlah ketidaksamaan objek data[22][23]. Proses K-medoids dimulai dengan memilih titik awal K. Titik ini adalah medoid yang berfungsi sebagai kandidat poin paling sentral dari kelompok tersebut. Kemudian setiap objek yang tersisa dikelompokkan dengan objek representatif yang paling mirip. Lalu mengganti salah satu medoid dengan medoid lain yang secara kualitas lebih baik dan dilakukan secara iterative.

Algoritma *Fuzzy c-means* (FCM) adalah algoritma pengelompokan khas yang memungkinkan titik data tertentu untuk berada dalam satu atau lebih *cluster*[21], dan tingkat keberadaan data dalam suatu *cluster* ditentukan oleh derajat keanggotaannya[24]. Tahapan kerja dari *fuzzy c-means* dimulai dengan menentukan banyak *cluster* (k) yang akan dibuat. Lalu dilanjutkan dengan menentukan nilai proporsi untuk setiap data poin secara random dengan persampaan (1), untuk masuk dalam suatu *cluster*.

$$\sum_{k=1}^c \mu_{ik} = 1 \quad (1)$$

dengan,

$c$  = jumlah cluster

$k$  = indeks cluster

$\mu_{ik}$  = nilai keanggotaan

Kemudian menghitung nilai centroid menggunakan persamaan (2).

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

dengan,

$V_{kj}$  = hasil pusat cluster yang dihasilkan

$n$  = jumlah data

$\mu_{ik}$  = nilai keanggotaan

$w$  = pangkat

$X_{ij}$  = data ke-i dan parameter ke-j

Kemudian menghitung kembali nilai proporsi untuk setiap data poin untuk masuk pada setiap cluster dengan persamaan (3).

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left( \left[ \sum_{j=1}^m (X_{ik} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (3)$$

### 2.3. Davies Bouldine Index (DBI)

Evaluasi internal pada model *clustering* dapat menggunakan *Davies Bouldine Index* (DBI), untuk menentukan jumlah *cluster* apakah sudah ideal atau belum[20]. Prinsip pendekatan DBI ini adalah memaksimalkan jarak antar *cluster* dan meminimalkan jarak didalam *cluster*[25].

### 2.5. Silhouette Coefficient

Evaluasi internal untuk hasil *clustering* dapat dilihat dari nilai *silhouette coefficient* yang menggunakan rumus jarak *Euclidean Distance*. Untuk nilai *silhouette coefficient* yang bagus adalah jika nilainya positif, yang artinya titik sudah berada pada *cluster* yang tepat. Jika nilai *silhouette coefficient* negatif artinya adanya *overlapping* sehingga titik berada pada *cluster* yang tidak tepat[26].

### 2.4. Proses Asosiasi

Pada pekerjaan *data mining*, asosiasi digunakan untuk mendapatkan aturan asosiasi dari pasangan item dengan menggunakan parameter *support* dan *confidence*. Nilai *support* adalah jumlah kemunculan pasangan item pada dataset, sedangkan *confidence* adalah kuatnya hubungan item dalam aturan asosiasi[14].

Algoritma yang digunakan dalam proses asosiasi ini adalah FP-Growth yang memiliki kinerja lebih cepat dibandingkan algoritma *a priori*, karena menghilangkan fase pembangkitan kandidat, menggantinya dengan konsep pembangunan pohon yang disebut FP-tree. Algoritma ini memiliki tahapan sebagai berikut:

1. Membuat *list item* dari setiap data transaksi;
2. Menghitung *support count* atau kemunculan item pada seluruh data transaksi;
3. Mengurutkan *item-set* berdasarkan nilai *support*;  $Support = P(A \cap B) = \frac{\Sigma \text{transaksi } A \text{ dan } B}{\Sigma \text{seluruh transaksi}}$  (4)
4. Membuat FP-tree;
5. Membuat *conditional pattern base* dari tiap *node* pada *tree*, yang berisi *parent* dari tiap *node* dan jumlah *support count* dari *node* yang merupakan *children*;
6. Membentuk *conditional FP-tree* yang diambil dari setiap *parent* pada *conditional pattern base*;
7. Membuat *frequent pattern*, yaitu kombinasi yang mungkin antar tiap *conditional FP-tree* dengan *node children*;
8. Filter *frequent pattern* yang sesuai dengan minimum *support* yang ditentukan.

Untuk menguji validitas dari aturan asosiasi yang terbentuk maka dapat digunakan *lift ratio* yang dihitung dengan persamaan (5).

$$lift = \frac{confidence}{jumlah transaksi} \quad (5)$$

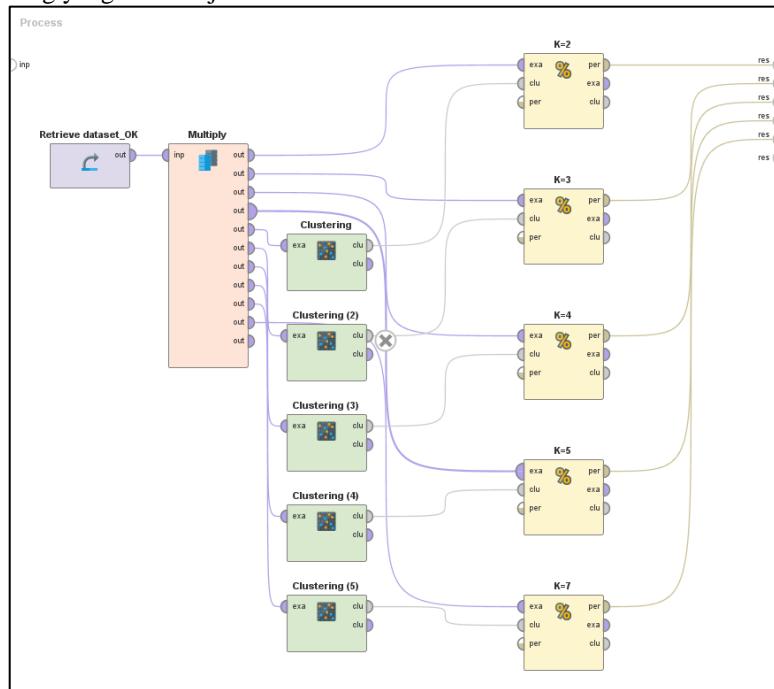
*Lift* menunjukkan adanya tingkat kekuatan *rule* atas kejadian acak dari *premise* dan *conclusion* berdasarkan nilai *support* masing-masing, dimana jika nilainya lebih dari 1 menunjukkan adanya manfaat dari aturan tersebut, serta semakin tinggi nilai *lift ratio* maka semakin besar kekuatan asosiasinya[27].

## 3. HASIL DAN PEMBAHASAN

Pada penelitian ini digunakan dataset berupa data transaksi penjualan makanan dan minuman

sejumlah 4.812 record yang diambil dari database sistem informasi penjualan. Data transaksi penjualan ini terdiri dari 2 tabel transaksi yang berrelasi dan memiliki 9 atribut yaitu id\_transaksi, tanggal & waktu transaksi, id\_detail\_transaksi, id\_item, nama item, harga, jumlah, sub total, total.

Masuk tahapan *preparation*, dataset tersebut dirubah dalam bentuk data transaksi seperti pada gambar 2. Transformasi data tersebut dilakukan untuk kebutuhan input pemodelan asosiasi menggunakan *FP-Growth*. Dataset terdiri dari atribut id\_transaksi yang membedakan antar transaksi penjualan yang terjadi, dan 42 item barang makanan dan minuman yang dijual. Setiap transaksi akan memiliki nilai 1 pada item barang yang terjual dan nilai 0 untuk item barang yang tidak terjual.



Gambar 3. Model clustering menggunakan *k-means* (M.C1)

Model M.C1 pada gambar 3 tersebut, setelah dijalankan menampilkan hasil evaluasi yang dapat dilihat pada tabel 1, dimana terdapat nilai DBI yang terkecil mendekati 0 tapi bukan negatif, yaitu 0.035 untuk  $k = 3$ .

Tabel 1. Nilai DBI untuk evaluasi M.C1

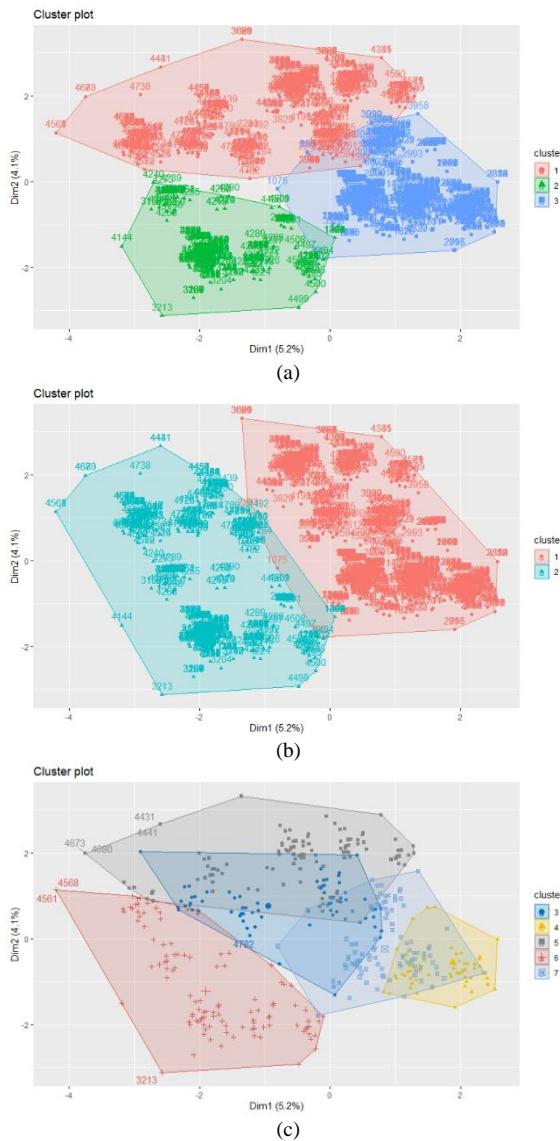
$k$	DBI		Silhouette Coefficient Fuzzy C-means
	K-means	K-medoids	
2	0.039	0.049	0.4043670
3	0.035	0.067	0.4989583
4	0.037	0.061	0.4852737
5	0.037	0.062	0.4975774
6	0.036	0.063	0.5457637
7	0.037	0.064	0.6215392
8	0.033	0.058	0.6168899
9	0.034	0.058	0.5692680

Untuk model clustering yang kedua (M.C2) menggunakan algoritma *k-medoids* dengan parameter yang sama dengan model M.C1, dan dihasilkan nilai DBI yang terkecil pada  $k=2$ , seperti yang terlihat pada tabel 1. Sedangkan untuk model

Tahapan selanjutnya *modeling* pertama, dilakukan *clustering* dengan menggunakan 3 algoritma *clustering* berbasis centroid, yaitu *k-means*, *k-medoids* dan *fuzzy c-means*. Pada proses *clustering* ini, penelitian berfokus untuk mencari  $k$  jumlah *cluster* yang baik untuk setiap algoritma *clustering*. Ujicoba dilakukan dengan nilai  $k$  yaitu 2, 3, 4, 5 dan 7.

Model *clustering* pertama (M.C1) pada gambar 3, disana model menggunakan algoritma *k-means* dengan parameter *euclidian distance* dan nilai  $k$  yaitu 2, 3, 4, 5 dan 7. Kemudian dilakukan evaluasi internal (*cluster distance performance*) menggunakan DBI.

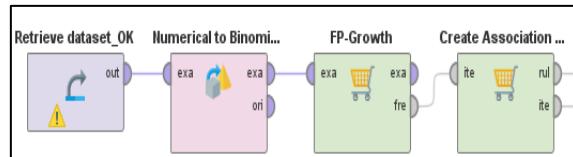
*clustering* ke-3 (M.C3) menggunakan algoritma *fuzzy c-means* dievaluasi menggunakan *silhouette coefficient*. Dari data pada tabel 1, didapatkan model *clustering* yang bagus adalah *clustering* dengan  $k=7$ . Untuk proses tahap pertama didapatkan model *clustering* yang baik berdasarkan algoritma masing-masing seperti pada Gambar 4 (a) untuk model K-Means; (b) untuk model K-Medoids; dan (c) untuk model Fuzzy c-means, dibawah ini.



Gambar 4. Visualisasi hasil Clustering masing-masing model MC.1, MC.2 dan M.C3

Tahapan pemodelan kedua adalah asosiasi untuk mendapatkan *frequent itemsets* dan membangkitkan *assosiation rules* menggunakan algoritma FP-Growth. Untuk tahapan ini dilakukan proses asosiasi menggunakan dataset dari tiap *cluster* yang terbentuk di tahapan *clustering* sebelumnya, dimana M.C1 menghasilkan 3 *cluster*, M.C2 menghasilkan 2 *cluster* dan M.C3 menghasilkan 7 *cluster* seperti pada gambar 4 diatas.

Pemodelan asosiasi yang pertama (M.A1) dengan dataset tanpa proses *clustering* sejumlah 4.812 baris data, menggunakan algoritma FP-Growth dengan parameter minimal *support* = 0.2 dan minimal *confidence* = 0.5 seperti pada gambar 5 dibawah ini.



Gambar 5. M.A1: Pemodelan FP-Growth pada dataset tanpa proses *clustering*

Dari proses asosiasi diatas dihasilkan 8 *frequent itemsets* yang berukuran 1 item dan 2 item saja, seperti pada tabel 2 dibawah ini. Dengan nilai minimal support = 0,2, maka *itemset* yang terpilih minimal kemunculan dalam dataset adalah 337 kali dari 4812 baris data.

Tabel 2. *frequent itemsets* dari M.A1

Size	Support ↗	Item 1	Item 2
1	0.42600	Indomie Rebus Telur,	
1	0.34400	Cireng Chili Oil,	
1	0.30800	Pisang Keju,	
1	0.30500	Vietnam drip,	
1	0.27000	Kopi Susu Regal,	
1	0.24800	Chocolate,	
2	0.25700	Pisang Keju,	Kopi Susu Regal,
2	0.25600	Cireng Chili Oil,	Vietnam drip,

Dari itemset diatas, kemudian dibuatkan *assosiation rules* dengan bentuk *premise* → *conclusion*, jika konsumen membeli item A (*premise*), maka dia akan membeli item B (*conclusion*) juga. Pembuatan *assosiation rules* dengan minimal *confidence* = 0.5, berarti jumlah kemunculan item *premise* dan item *conclusion* secara bersamaan minimalnya 50% dari jumlah kemunculan item *premise*.

Tabel 3. *Association Rules* dari M.A1

Premise	Conclusion	Support	Confidence	Lift
Pisang Keju,	Kopi Susu Regal,	0.257481297	0.835468645	3.092519321
Kopi Susu Regal,	Pisang Keju,	0.257481297	0.953076923	3.092519321
Cireng Chili Oil,	Vietnam drip,	0.255818786	0.743357488	2.438334173
Vietnam drip.	Cirene Chili Oil.	0.255818786	0.839127471	2.438334173

Pada tabel 3 diatas, diketahui hasil pembangkitan *assosiation rules* dari M.A1 didapatkan 4 aturan yang memenuhi minimal *support* dan *confidence*. Berdasarkan nilai *Lift*-nya, 4 aturan tersebut dapat dikatakan valid karena memiliki nilai lebih dari 1 yang menunjukkan adanya manfaat dari aturan tersebut.

Pemodelan asosiasi yang kedua (M.A2) menggunakan dataset hasil *clustering* dengan algoritma *k-means*, dimana dataset dikelompokkan dalam 3 *cluster* masing-masing berisi 1.336, 1.712 dan 1.764 baris data. M.A2 menggunakan algoritma FP-Growth dengan parameter yang sama dengan M.A1, yaitu minimal *support* = 0.2 dan minimal *confidence* = 0.5.

Tabel 4. frequent itemsets dari M.A2

C	Size	Support	Item 1	Item 2	Item 3	Item 4
0	1	0.999	Pisang Keju,			
0	1	0.928	Kopi Susu Regal,			
0	1	0.313	Vietnam drip,			
0	1	0.299	Cireng Chili Oil,			
0	2	0.927	Pisang Keju,	Kopi Susu Regal,		
0	2	0.313	Pisang Keju,	Vietnam drip,		
0	2	0.312	Kopi Susu Regal,	Vietnam drip,		
0	2	0.299	Pisang Keju,	Cireng Chili Oil,		
0	2	0.299	Kopi Susu Regal,	Cireng Chili Oil,		
0	2	0.264	Vietnam drip,	Cireng Chili Oil,		
0	3	0.312	Pisang Keju,	Kopi Susu Regal,	Vietnam drip,	
0	3	0.299	Pisang Keju,	Kopi Susu Regal,	Cireng Chili Oil,	
0	3	0.264	Pisang Keju,	Vietnam drip,	Cireng Chili Oil,	
0	3	0.264	Kopi Susu Regal,	Vietnam drip,	Cireng Chili Oil,	
0	4	0.264	Pisang Keju,	Kopi Susu Regal,	Vietnam drip,	Cireng Chili Oil,
1	1	0.542	Cireng Chili Oil,			
1	1	0.524	Vietnam drip,			
1	2	0.464	Cireng Chili Oil,	Vietnam drip,		
2	1	0.928	Indomie Rebus Telur,			
2	1	0.546	Chocolate,			
2	2	0.474	Indomie Rebus Telur,	Chocolate,		

Model M.A2 ini menghasilkan 21 *itemset* prioritas dari masing-masing *cluster* seperti pada tabel 4, terdiri dari 1-*itemset* sampai 4-*itemset* yang memiliki support minimal 0.2. Dataset *cluster-0* menghasilkan *itemset* yang lebih banyak dibanding *cluster* lainnya. Sedangkan untuk aturan asosiasinya dihasilkan 34 aturan asosiasi yang memenuhi minimal *support* dan *confidence*. Seperti yang terlihat pada tabel 5, 34 aturan asosiasi dari M.A2 ini sudah memiliki nilai *lift* yang bagus yaitu lebih besar sama dengan 1.

Tabel 5. Association Rules dari M.A2

C	Premise	Conclusion	Support	Confidence	Lift
0	Vietnam drip,	Pisang Keju,	0.31287	1.00000	1.00075
0	Kopi Susu Regal,, Vietnam drip,	Pisang Keju,	0.31213	1.00000	1.00075
0	Vietnam drip,	Pisang Keju,, Kopi Susu Regal,	0.31213	0.99761	1.07571
0	Pisang Keju,, Vietnam drip,	Kopi Susu Regal,	0.31213	0.99761	1.07484
0	Vietnam drip,	Kopi Susu Regal,	0.31213	0.99761	1.07484
0	Cireng Chili Oil,,	Pisang Keju,, Kopi Susu Regal,	0.29940	1.00000	1.07829
0	Cireng Chili Oil,,	Kopi Susu Regal,,	0.29940	1.00000	1.07742
0	Pisang Keju,, Cireng Chili Oil,,	Kopi Susu Regal,,	0.29940	1.00000	1.07742
0	Cireng Chili Oil,,	Pisang Keju,,	0.29940	1.00000	1.00075
0	Kopi Susu Regal,, Cireng Chili Oil,,	Pisang Keju,,	0.29940	1.00000	1.00075
0	Vietnam drip,, Cireng Chili Oil,,	Kopi Susu Regal,,	0.26422	1.00000	1.07829
0	Pisang Keju,, Vietnam drip,, Cireng Chili Oil,,	Kopi Susu Regal,,	0.26422	1.00000	1.07742
0	Vietnam drip,, Cireng Chili Oil,,	Kopi Susu Regal,, Pisang Keju,,	0.26422	1.00000	1.00075
0	Kopi Susu Regal,, Vietnam drip,, Cireng Chili Oil,,	Pisang Keju,,	0.26422	1.00000	1.00075
0	Vietnam drip,, Cireng Chili Oil,,	Kopi Susu Regal,, Pisang Keju,,	0.26422	0.88250	2.82739
0	Cireng Chili Oil,,	Pisang Keju,, Kopi Susu Regal,, Vietnam drip,,	0.26422	0.88250	2.82739
0	Pisang Keju,, Cireng Chili Oil,,	Kopi Susu Regal,, Vietnam drip,,	0.26422	0.88250	2.82739
0	Pisang Keju,, Kopi Susu Regal,, Cireng Chili Oil,,	Vietnam drip,,	0.26422	0.88250	2.82739
0	Cireng Chili Oil,,	Vietnam drip,, Pisang Keju,,	0.26422	0.88250	2.82739
0	Pisang Keju,, Pisang Keju,, Vietnam drip,,	Vietnam drip,,	0.26422	0.88250	2.82062
0	Kopi Susu Regal,, Cireng Chili Oil,,	Vietnam drip,,	0.26422	0.88250	2.82062
0	Pisang Keju,, Cireng Chili Oil,,	Vietnam drip,,	0.26422	0.88250	2.82062
0	Pisang Keju,, Kopi Susu Regal,, Cireng Chili Oil,,	Vietnam drip,,	0.26422	0.88250	2.82062
0	Kopi Susu Regal,, Vietnam drip,,	Pisang Keju,,	0.26422	0.84652	2.82739
0	Kopi Susu Regal,, Vietnam drip,,	Pisang Keju,, Cireng Chili Oil,,	0.26422	0.84652	2.82739
0	Pisang Keju,, Kopi Susu Regal,, Vietnam drip,,	Cireng Chili Oil,,	0.26422	0.84652	2.82739
0	Pisang Keju,, Vietnam drip,,	Cireng Chili Oil,,	0.26422	0.84590	2.82062
0	Pisang Keju,, Kopi Susu Regal,, Cireng Chili Oil,,	Vietnam drip,,	0.26422	0.84590	2.82062
0	Vietnam drip,,	Pisang Keju,, Cireng Chili Oil,,	0.26422	0.84590	2.82062
0	Vietnam drip,,	Pisang Keju,, Kopi Susu Regal,, Cireng Chili Oil,,	0.26422	0.84590	2.82062
1	Vietnam drip,,	Cireng Chili Oil,,	0.46437	0.88629	1.63505

Pemodelan asosiasi yang ketiga (M.A3) menggunakan dataset hasil *clustering* dengan algoritma *k-medoids*, dimana dataset dikelompokan dalam 2 *cluster* masing-masing berisi 2.046 dan 2.767 baris data. M.A3 menggunakan algoritma FP-Growth dengan parameter yang sama dengan M.A1, yaitu minimal *support* = 0.2 dan minimal *confidence* = 0.5.

Tabel 6. frequent itemsets dari M.A3

C	Size	Support	Item 1	Item 2	Item 3
0	1	0.807	Cireng Chili Oil,,		
0	1	0.717	Vietnam drip,,		
0	1	0.377	Indomie Rebus Telur,,		
0	1	0.288	Pisang Keju,,		
0	1	0.272	Kopi Susu Regal,,		
0	2	0.602	Cireng Chili Oil,,	Vietnam drip,,	
0	2	0.317	Cireng Chili Oil,,	Indomie Rebus Telur,,	
0	2	0.247	Pisang Keju,,	Kopi Susu Regal,,	
0	2	0.237	Cireng Chili Oil,,	Pisang Keju,,	
0	2	0.237	Vietnam drip,,	Pisang Keju,,	
0	2	0.233	Vietnam drip,,	Indomie Rebus Telur,,	
0	2	0.229	Vietnam drip,,	Kopi Susu Regal,,	
0	2	0.221	Cireng Chili Oil,,	Kopi Susu Regal,,	
0	3	0.206	Cireng Chili Oil,,	Vietnam drip,,	Pisang Keju,,
0	3	0.203	Vietnam drip,,	Pisang Keju,,	Kopi Susu Regal,,
1	1	0.463	Indomie Rebus Telur,,		
1	1	0.323	Pisang Keju,,		
1	1	0.307	Chocolate,,		
1	1	0.269	Kopi Susu Regal,,		
1	2	0.265	Pisang Keju,,	Kopi Susu Regal,,	
1	2	0.242	Indomie Rebus Telur,,	Chocolate,,	

Model M.A2 ini menghasilkan 12 *itemset* prioritas dari masing-masing *cluster* seperti pada tabel 6, terdiri dari 1-*itemset* sampai 3-*itemset* yang memiliki support minimal 0.2. Dataset *cluster-0* menghasilkan *itemset* yang lebih banyak dibanding *cluster* lainnya. Sedangkan untuk aturan asosiasinya dihasilkan 34 aturan asosiasi yang memenuhi minimal *support* dan *confidence*. Seperti yang terlihat pada tabel 6, 20 aturan asosiasi dari M.A3 ini sudah memiliki nilai *lift* yang bagus yaitu lebih besar sama dengan 1.

Tabel 7. Association Rules dari M.A2

C	Premise	Conclusion	Support	Confidence	Lift
0	Kopi Susu Regal,,	Pisang Keju,,	0.24682	0.90664	3.14405
0	Vietnam drip,, Kopi Susu Regal,,	Pisang Keju,,	0.20332	0.88889	3.08249
0	Cireng Chili Oil,, Pisang Keju,,	Vietnam drip,,	0.20577	0.86983	1.21397
0	Vietnam drip,, Pisang Keju,,	Cireng Chili Oil,,	0.20577	0.86804	1.07507
0	Vietnam drip,, Pisang Keju,,	Kopi Susu Regal,,	0.20332	0.85773	3.15066
0	Pisang Keju,,	Kopi Susu Regal,,	0.24682	0.85593	3.14405
0	Indomie Rebus Telur,,	Cireng Chili Oil,,	0.31672	0.84047	1.04092
0	Kopi Susu Regal,,	Vietnam drip,,	0.22874	0.84022	1.17263
0	Vietnam drip,,	Cireng Chili Oil,,	0.60166	0.83970	1.03997
0	Pisang Keju,, Kopi Susu Regal,,	Vietnam drip,,	0.20332	0.82376	1.14967
0	Pisang Keju,,	Vietnam drip,,	0.23705	0.82203	1.21726
0	Pisang Keju,,	Cireng Chili Oil,,	0.23656	0.82034	1.01599
0	Kopi Susu Regal,,	Cireng Chili Oil,,	0.22092	0.81149	1.00503
0	Kopi Susu Regal,,	Vietnam drip,, Pisang Keju,,	0.20332	0.74686	3.15066
0	Cireng Chili Oil,,	Vietnam drip,,	0.60166	0.74516	1.03997
0	Pisang Keju,,	Cireng Chili Oil,, Vietnam drip,,	0.20577	0.71356	1.18598
0	Pisang Keju,,	Vietnam drip,, Kopi Susu Regal,,	0.20332	0.70508	3.08249
1	Kopi Susu Regal,,	Pisang Keju,,	0.26537	0.98789	3.05991
1	Pisang Keju,,	Kopi Susu Regal,,	0.26537	0.82195	3.05991
1	Chocolate,,	Indomie Rebus Telur,,	0.24187	0.78892	1.70480

Pemodelan asosiasi yang terakhir (M.A4) menggunakan dataset hasil *clustering* dengan algoritma *fuzzy c-means*, dimana dataset dikelompokan dalam 7 *cluster*. M.A4 menggunakan algoritma FP-Growth dengan parameter yang sama dengan M.A1, yaitu minimal *support* = 0.2 dan minimal *confidence* = 0.5.

Model M.A4 ini menghasilkan 84 *itemset* prioritas dari masing-masing *cluster* yang cukup banyak, terdiri dari 1-*itemset* sampai 6-*itemset* yang memiliki support minimal 0.2. Dataset *cluster-1* menghasilkan *itemset* yang lebih banyak dibanding *cluster* lainnya. Sedangkan untuk aturan asosiasinya dihasilkan 4 aturan asosiasi yang memenuhi minimal *support* dan *confidence*. Seperti yang terlihat pada tabel 6, 243 aturan asosiasi dari M.A4 ini sudah

memiliki nilai *lift* yang bagus yaitu lebih besar sama dengan 1.

Tabel 6. Association Rules dari M.A4

C	Premise	Conclusion	Support	Confidence	Lift
1	Indomie Rebus Telur,	Cireng Chili Oil,, Pisang Keju,	0.7839%	1.00000	1.25532
1	Indomie Rebus Telur,	Pisang Keju,	0.7839%	1.00000	1.20408
1	Cireng Chili Oil,, Indomie Rebus Telur,	Pisang Keju,	0.7839%	1.00000	1.20408
1	Indomie Rebus Telur,	Cireng Chili Oil,	0.7839%	1.00000	1.11321
1	Pisang Keju,, Indomie Rebus Telur,	Cireng Chili Oil,	0.7839%	1.00000	1.11321
1	Cireng Chili Oil,, Vietnam drip,	Indomie Rebus Telur,	0.72034	1.00000	1.27568
1	Cireng Chili Oil,, Vietnam drip,	Pisang Keju,, Indomie Rebus Telur,	0.72034	1.00000	1.27568
1	Cireng Chili Oil,, Pisang Keju,, Vietnam drip,	Indomie Rebus Telur,	0.72034	1.00000	1.27568
1	Vietnam drip,, Indomie Rebus Telur,	Cireng Chili Oil,, Pisang Keju,	0.72034	1.00000	1.25532
1	Cireng Chili Oil,, Vietnam drip,	Pisang Keju,	0.72034	1.00000	1.20408
1	Vietnam drip,, Indomie Rebus Telur,	Pisang Keju,	0.72034	1.00000	1.20408
1	Cireng Chili Oil,, Vietnam drip,, Indomie Rebus Telur,	Pisang Keju,	0.72034	1.00000	1.20408
1	Vietnam drip,, Indomie Rebus Telur,	Cireng Chili Oil,	0.72034	1.00000	1.11321
1	Pisang Keju,, Vietnam drip,, Indomie Rebus Telur,	Cireng Chili Oil,	0.72034	1.00000	1.11321
1	Kopi Susu Regal,,	Indomie Rebus Telur,	0.59746	1.00000	1.27568
1	Kopi Susu Regal,,	Cireng Chili Oil,, Indomie Rebus Telur,	0.59746	1.00000	1.27568
1	Cireng Chili Oil,, Kopi Susu Regal,,	Indomie Rebus Telur,	0.59746	1.00000	1.27568

#### 4. DISKUSI

Maka dari hasil penelitian ini didapatkan rekap jumlah aturan asosiasi yang dihasilkan pada model asosiasi menggunakan FP-Growth seperti pada tabel 7 dibawah ini.

Tabel 6. Perbandingan hasil asosiasi FP-Growth

Asosiasi	Dataset	Clustering	Jumlah cluster	min. support	jumlah frequent itemset	min. confidence	jumlah aturan asosiasi	rata-rata lift ratio
FP-Growth	dataset_ALL	-	1		9		4	3.496830572
		K-Medoids	2	0.2	12	0.5	20	2.339046133
	dataset_clustered	K-Means	3		21		34	3.185168611
		Fuzzy C-Means	7		84		243	1.204623843

Seluruh model asosiasi yang dibangun mulai dari M.A1 sampai M.A4, menggunakan nilai parameter yang sama yaitu *min. Support* = 0.2 dan *min. Confidence* = 0.5, hal ini disesuaikan dengan kemunculan aturan asosiasi pada model M.A1 yang menggunakan dataset ukuran besar. Aturan asosiasi baru dihasilkan saat ambang batas *support* berada dinilai 0.2 dan *confidence* 0.5.

Model asosiasi M.A2 sampai dengan M.A4 menggunakan dataset yang sudah diklasterisasi berdasarkan nilai kemiripan /jarak. Walaupun menggunakan nilai parameter yang sama, namun dengan adanya pengelompokan dataset berdasarkan kemiripannya pada proses *clustering*, jumlah *frequent itemset* dan aturan asosiasi menjadi bertambah banyak. Hal ini karena dataset yang diklasterisasi mengurangi ukuran dari dataset yang digunakan untuk membangun aturan asosiasi. Sehingga dengan nilai parameter yang sama, sudah dapat menghasilkan jumlah *frequent itemset* dan aturan asosiasi yang lebih banyak.

Dari model M.A2, M.A3 dan M.A4, dapat dilihat rata-rata *lift ratio* yang tertinggi ada di M.A2 yang menggunakan dataset hasil klasterisasi algoritma *k-means*. Hal ini menunjukan *kmeans* mampu mengelompokan data yang memiliki kemiripan tinggi pada setiap *cluster*-nya, sehingga pada proses asosiasi mempermudah pencarian *frequent itemset* dan aturan asosiasi.

Dari penelitian sebelumnya[12], penelitian ini memiliki 2 algoritma *clustering* yang berbeda namun memiliki karakter yang sama dengan *k-medoids* yaitu *k-means* dan *fuzzy c-means*. Dan didapatkan model M.A2 memiliki rata-rata *lift ratio* tertinggi yaitu model yang menggunakan *k-means* untuk proses *clustering* dataset kedalam 3 *cluster* kemudian tiap *cluster* diproses menggunakan algoritma *Fp-Growth*.

Dari model M.A2 ini, didapatkan 32 aturan asosiasi yang unik dari 34 aturan asosiasi yang dihasilkan. Daftar pasangan item yang sering terjual bersamaan dapat dilihat pada gambar 6 dibawah ini.

IF Cireng Chili Oil, THEN Kopi Susu Regal,
IF Cireng Chili Oil, THEN Kopi Susu Regal,, Vietnam drip,
IF Cireng Chili Oil, THEN Pisang Keju,
IF Cireng Chili Oil, THEN Pisang Keju,, Kopi Susu Regal,
IF Cireng Chili Oil, THEN Pisang Keju,, Kopi Susu Regal,, Vietnam drip,
IF Cireng Chili Oil, THEN Pisang Keju,, Vietnam drip,
IF Cireng Chili Oil, THEN Vietnam drip,
IF Kopi Susu Regal,, Cireng Chili Oil, THEN Pisang Keju,
IF Kopi Susu Regal,, Cireng Chili Oil, THEN Pisang Keju,, Vietnam drip,
IF Kopi Susu Regal,, Cireng Chili Oil, THEN Vietnam drip,
IF Kopi Susu Regal,, Vietnam drip, THEN Pisang Keju,
IF Kopi Susu Regal,, Vietnam drip, THEN Pisang Keju,, Cireng Chili Oil,
IF Kopi Susu Regal,, Vietnam drip,, Cireng Chili Oil, THEN Pisang Keju,
IF Pisang Keju,, Cireng Chili Oil, THEN Kopi Susu Regal,
IF Pisang Keju,, Cireng Chili Oil, THEN Kopi Susu Regal,, Vietnam drip,
IF Pisang Keju,, Cireng Chili Oil, THEN Vietnam drip,
IF Pisang Keju,, Kopi Susu Regal,, Cireng Chili Oil, THEN Vietnam drip,
IF Pisang Keju,, Kopi Susu Regal,, Vietnam drip, THEN Cireng Chili Oil,
IF Pisang Keju,, Vietnam drip, THEN Cireng Chili Oil,
IF Pisang Keju,, Vietnam drip, THEN Kopi Susu Regal,
IF Pisang Keju,, Vietnam drip, THEN Kopi Susu Regal,, Cireng Chili Oil,
IF Pisang Keju,, Vietnam drip,, Cireng Chili Oil, THEN Kopi Susu Regal,
IF Vietnam drip, THEN Cireng Chili Oil,
IF Vietnam drip, THEN Kopi Susu Regal,
IF Vietnam drip, THEN Kopi Susu Regal,, Cireng Chili Oil,
IF Vietnam drip, THEN Pisang Keju,
IF Vietnam drip, THEN Pisang Keju,, Cireng Chili Oil,
IF Vietnam drip, THEN Pisang Keju,, Kopi Susu Regal,
IF Vietnam drip, THEN Pisang Keju,, Kopi Susu Regal,, Cireng Chili Oil,
IF Vietnam drip,, Cireng Chili Oil, THEN Kopi Susu Regal,
IF Vietnam drip,, Cireng Chili Oil, THEN Pisang Keju,
IF Vietnam drip,, Cireng Chili Oil, THEN Pisang Keju,, Kopi Susu Regal,

Gambar 5. Pasangan item yang sering terjual bersamaan berdasarkan Model M.A2

Dari daftar pasangan item di Gambar 5 diatas, pemilik memiliki rekomendasi untuk melakukan strategi penjualan agar dapat meningkatkan profitnya. Salah satunya, adalah menyiapkan stok yang banyak untuk item cireng chili oil, kopi susu regal, pisang keju dan vietnam drip. Karena berdasarkan gambar 5 baris pertama, jika pembeli memesan Cireng Chili Oil maka akan memesan Kopi Susu Regal.

Berdasarkan daftar aturan asosiasi pada Gambar 5 juga, dapat dilakukan promosi paket menu yang direkomendasikan untuk pembeli baru seperti jika membeli Pisang Keju dan Vietnam Drip, maka dapat direkomendasikan Kopi Susu Regal dengan Cireng Chili Oil.

## 5. KESIMPULAN

Dari penelitian ini didapatkan hasil dari pencarian jumlah cluster yang ideal pada proses clustering masing-masing algoritma berdasarkan nilai DBI dan *Silhouette index*, yaitu  $k=3$  untuk *kmeans*,  $k=2$  untuk *kmedoids* dan  $k=7$  untuk *fuzzy c-means*. Kemudian hasil pencari aturan asosiasi dari tiap cluster dengan nilai lift ratio yang baik yaitu diatas 1 seperti pada tabel 7. Selain dari itu, hasil pemodelan algoritma centroid clustering *fuzzy c-means* dapat menghasilkan jumlah aturan asosiasi dengan yang terbanyak, namun model yang menggunakan *kmeans* menghasilkan aturan asosiasi dengan rata-rata *lift ratio* tertinggi.

Hasil dari model asosiasi menggunakan *k-means* dan *FP-Growth* menghasilkan aturan asosiasi yang disusun dari item yang sering muncul adalah cireng chili oil, kopi susu regal, pisang keju dan vietnam drip. Sehingga pemilik toko dapat menyiapkan stok yang cukup banyak untuk 4 item tersebut sehingga permintaan pelanggan selalu terpenuhi.

Untuk pengembangan penelitian ini kedepannya, disarankan untuk menambahkan ukuran dataset serta menggunakan algoritma *clustering* yang bersifat *soft-clustering* agar dapat memperbaiki pengelompokan dataset sebelum proses asosiasi.

## DAFTAR PUSTAKA

- [1] S. Rijal, M. I. S. Ahmad, N. Arisah, Nurdiana, and M. Hasan, "Kajian Peluang dan Tantangan UMKM: Sanggupkah Sektor Kuliner Bertahan di Masa Pandemi COVID-19," *Proceeding Teknol. Pendidik. Semin. Daring Nas. 2021 Digit. Gener. Digit. Nation*, pp. 115–1225, 2021.
- [2] Taufikurrahman, "Analisis penjualan online ditengah pandemi covid-19 Analysis of online sales in the midst of the covid-19 pandemic," *J. Ekon. dan Manaj.*, vol. 18, no. 3, pp. 440–445, 2021.
- [3] S. N. Rohmah, "Adakah peluang bisnis di tengah kelesuan perekonomian akibat pandemi Corona virus Covid-19?," *'ADALAH Bul. Huk. Keadilan*, vol. 4, no. 1, pp. 63–74, 2020.
- [4] D. Rusdianto, Sutiyono, and L. Zaelani, "Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Mengetahui Pola Peminjaman Buku Di Perpustakaan Universitas Bale Bandung," *J-Sika*, vol. 02, no. 02, pp. 1–10, 2020, [Online]. Available: <https://ejournal.uniibba.ac.id/index.php/j-sika/article/view/376/313>.
- [5] K. N. Wijaya, R. F. Malik, and S. Nurmaini, "Analisa Pola Frekuensi Keranjang Belanja Dengan Dengan Perbandingan Algoritma Fp-Growth (Frequent Pattern Growth) dan Eclat pada minimarket," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 2, pp. 364–373, 2020, doi: 10.35957/jatisi.v7i2.380.
- [6] L. Indah Prahartiwi, S. Informasi, S. Nusa Mandiri, J. Damai No, and W. Jati Barat Jakarta Selatan DKI Jakarta, "Pencarian Frequent Itemset pada Analisis Keranjang Belanja Menggunakan Algoritma FP-Growth," *Inf. Syst. Educ. Prof.*, vol. 2, no. 1, pp. 1–10, 2017.
- [7] R. Sari and R. Y. Hayuningtyas, "Analisis Keranjang Belanja Pada Transaksi Penjualan Menggunakan Algoritma Apriori," *EVOLUSI J. Sains dan Manaj.*, vol. 9, no. 1, pp. 46–51, 2021, doi: 10.31294/evolusi.v9i1.9999.
- [8] N. R. S. Purba and F. Riandari, "Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Analisis Keranjang Belanja Pada Transaksi Penjualan Pada PT Madu Kembang Joyo," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 4, no. 1, pp. 69–74, 2021, doi: 10.32672/jnkti.v4i1.2745.
- [9] M. S. B. Butar and E. Elisa, "RULES ASSOCIATION FP-GROWTH DALAM ANALISIS KERANJANG PASAR," *Comasie*, vol. 6, no. 2, pp. 127–136, 2022.
- [10] D. Cahyanti and I. Permana, "Comparison of Book Shopping Patterns Before and During the Covid-19 Pandemic Using the Fp-Growth Algorithm At Zanafa Perbandingan Pola Belanja Buku Sebelum Dan Masa Pandemi Covid-19 Menggunakan Algoritma Fp-Growth Pada Toko Buku," *J. Tek. Inform.*, vol. 3, no. 2, pp. 381–386, 2022.
- [11] S. Wahyuni, Suherman, and lumalo portibi Harahap, "Implementasi Data Mining dalam Memprediksi Stok Barang Menggunakan Algoritma Apriori," *J. Tek. DAN Inform.*, vol. 5, no. 2, pp. 67–71, 2018.
- [12] M. I. Ghozali, R. Z. Ehwan, and W. H. Sugiharto, "Analisa Pola Belanja Menggunakan Algoritma Fp Growth, Self Organizing Map (Som) Dan K Medoids," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 8, no. 1, pp. 317–326, 2017, doi: 10.24176/simet.v8i1.995.
- [13] X. Wu *et al.*, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, 2008, doi: 10.1007/s10115-007-0114-2.
- [14] I. Syukra, A. Hidayat, and M. Z. Fauzi, "Implementation of K-Medoids and FP-Growth Algorithms for Grouping and Product Offering Recommendations," *Indones. J. Artif. Intell. Data Min.*, vol. 2, no. 2, p. 107, 2019, doi: 10.24014/ijaidm.v2i2.8326.

- [15] N. Mamahit and A. Qoiriah, "Penerapan Algoritma Fp-Growth dan K-Means pada Data Transaksi Minimarket," *J. Informatics Comput. Sci.*, vol. 1, no. 02, pp. 78–83, 2019.
- [16] N. P. Dharshinni, E. Bangun, S. Karunia, R. Damayanti, G. Rophe, and R. Pandapotan, "Menu Package Recommendation using Combination of K-Means and FP-Growth Algorithms at Bakery Stores," *J. Mantik*, vol. 4, no. 2, pp. 1272–1277, 2020.
- [17] C. Mustika, S. G. Setyorini, E. K. Sari, L. R. Elita, and S. A. Putri, "Market Basket Analysis with K-Means and FP-Growth Algorithm as Citra Mustika Pandawa Company Analisis Keranjang Pasar Menggunakan Algoritma K-Means dan," *Inst. Res. Publ. Indones.*, vol. 1, no. April, pp. 41–46, 2021.
- [18] E. Muningsih, I. Maryani, and V. R. Handayani, "Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa," *J. Sains dan Manaj.*, vol. 9, no. 1, pp. 95–100, 2021, [Online]. Available: <https://ejournal.bsi.ac.id/ejurnal/index.php/evalusi/article/view/10428/4839>.
- [19] I. L. Putri *et al.*, "Aturan asosiasi bahan pada resep jamu tradisional dengan algoritma apriori," in *Proceeding SENDIU 2020*, 2020, pp. 978–979.
- [20] A. Badruttamam, S. Sudarno, and D. A. I. Maruddani, "PENERAPAN ANALISIS KLASTER K-MODES DENGAN VALIDASI DAVIES BOULDIN INDEX DALAM MENENTUKAN KARAKTERISTIK KANAL YOUTUBE DI INDONESIA (Studi Kasus: 250 Kanal YouTube Indonesia Teratas Menurut Socialblade)," *J. Gaussian*, vol. 9, no. 3, pp. 263–272, 2020, doi: 10.14710/j.gauss.v9i3.28907.
- [21] S. K. Uppada, "Centroid Based Clustering Algorithms- A Clarion Study," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 6, pp. 7309–7313, 2014, [Online]. Available: <https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.658.3904>.
- [22] R. D. Ramadhan and D. J. Ak, "Evaluasi K-Means dan K-Medoids pada Dataset Kecil," *Semin. Nas. Inform. dan Apl.*, no. September, pp. 20–24, 2017, [Online]. Available: [file:///C:/Users/User/Downloads/Evaluasi K-Means dan K-Medoids pada Dataset Kecil.pdf](file:///C:/Users/User/Downloads/Evaluasi%20K-Means%20dan%20K-Medoids%20pada%20Dataset%20Kecil.pdf).
- [23] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. I. R.H.Zer, and D. Hartama, "Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 166–173, 2020, doi: 10.36294/jurti.v4i1.1296.
- [24] A. Ramadhan, Mustakim, and R. Handinata, "Implementasi Algoritma Fuzzy C Means Dan Moora Untuk Pengelompokan Dan Penentuan Wilayah Penanggulangan Bencana Banjir," no. November, p. Pekanbaru, 2019.
- [25] W. Gie and D. Jollyta, "Perbandingan Euclidean dan Manhattan Untuk Optimasi Cluster Menggunakan Davies Bouldin Index : Status Covid-19 Wilayah Riau," *Pros. Semin. Nas. Ris. Dan Inf. Sci. 2020*, vol. 2, no. April, pp. 187–191, 2020.
- [26] E. B. Susanto, "Evaluasi Hasil Klaster Pada Dataset Iris , Soybean-small , Wine Menggunakan Algoritma Fuzzy C-Means dan K-," *Surya Inform.*, vol. 2, no. 1, pp. 6–13, 2016.
- [27] Despitaria, H. Sujaini, and Tursina, "Analisis Asosiasi pada Transaksi Obat Menggunakan Data Mining dengan Algoritma A Priori," *Justin*, vol. 4, no. 2, p. 6, 2016.