

OPTIMIZATION PRODUCT RECOMMENDATION USING K-MEANS, AGGLOMERATIVE CLUSTERING AND FP-GROWTH ALGORITHM

Ratu Najmil Huda^{*1}, Rifqi Fitriadi², Arief Wibowo³

^{1,2}Computer Science Master's Study Program, Faculty of Information Technology, Universitas Budi Luhur, Indonesia

³Information Systems Study Program, Faculty of Information Technology, Universitas Budi Luhur, Indonesia
Email: ratu.najmilhuda@gmail.com, rifqi0587@gmail.com

(Article received: February 27, 2024; Revision: April 12, 2024; published: July 29, 2024)

Abstract

The growth of online business has been rising considerably in recent years. The growth is affected by technology advancement in Internet and smartphones and consumer behavior change for better online shopping experience. To anticipate this swift customer behavior, business owners need to have an excellent inventory management to be able to keep making profits. In data mining realm, the algorithm model that is known to be applied in this case is the association algorithm. This model will explicate customers' purchasing patterns where is useful in calculating stock accurately. The aim of this research is to find an appropriate model in handling large data to obtain valid association rules that have minimum support value, confidence value, and high lift ratios. It is hoped that the results of this research can provide recommendations for online sellers to manage a large variety of goods and to keep making profits. Datasets that contain a large variety of goods are handled first by using a clustering algorithm to group similar items together. The dataset tested was divided into three groups, namely, dataset without clusters, k-means cluster, and agglomerative cluster. After forming three groups of datasets, FP-Growth was applied to each dataset. The result is that datasets with clusters, whether using k-means or agglomerative, have a minimum support value that is greater than datasets without clusters. Most association rules are obtained from the k-means cluster dataset. Based on the model applied in this research, the association itemset size only obtains one conclusion from one premise.

Keywords: Agglomerative Clustering, FP-Growth, K-Means, Product Recommendations.

OPTIMALISASI REKOMENDASI PRODUK MENGGUNAKAN ALGORITMA K-MEANS, AGGLOMERATIVE CLUSTERING, DAN FP-GROWTH

Abstrak

Perkembangan bisnis online meningkat pesat dalam beberapa tahun terakhir. Hal ini didorong oleh perkembangan teknologi internet dan smartphone serta perubahan perilaku konsumen yang menyukai kemudahan dan kenyamanan berbelanja. Penjualan makanan online merupakan salah satu jenis usaha yang menjanjikan bagi pemilik bisnis, agar selalu meningkatkan profit, salah satu strategi bisnis yang harus dimiliki jenis usaha ini adalah pengelolaan stok. Dalam dunia data mining, model algoritma yang sudah dikenal untuk diaplikasikan dalam kasus ini yaitu algoritma asosiasi. Algoritma ini membantu menginterpretasikan pola pembelian pelanggan, berguna untuk menentukan perhitungan stok yang akurat. Tujuan dari penelitian ini adalah menemukan model yang tepat dalam penanganan data yang besar untuk mendapatkan aturan asosiasi dengan nilai minimal *support*, nilai *confidence* dan *lift ratio* yang tinggi dan valid. Dataset yang banyak variasi barangnya dilakukan penanganan terlebih dahulu menggunakan algoritma clustering guna mengelompokkan barang yang serupa. Dataset yang diuji dibagi dalam tiga kelompok yaitu, dataset tanpa cluster, dataset cluster k-means dan dataset cluster agglomerative. Setelah terbentuk tiga kelompok dataset, baru kemudian dilanjutkan pengaplikasian FP-Growth ke masing-masing dataset. Hasilnya *dataset* dengan *cluster*, baik menggunakan metode *k-means* maupun *agglomerative*, memiliki nilai minimal *support* yang lebih besar dibandingkan dengan *dataset* tanpa *cluster*. Aturan asosiasi paling banyak didapat dari *dataset cluster k-means*. Atas model yang diterapkan pada penelitian ini, *size itemset* asosiasi hanya didapat satu konklusi dari satu *premise*.

Kata kunci: Agglomerative Clustering, FP-Growth, K-Means, Rekomendasi Produk.

1. PENDAHULUAN

Perkembangan bisnis online di Indonesia telah meningkat pesat dalam beberapa tahun terakhir [1]. Hal ini didorong oleh perkembangan teknologi internet dan smartphone yang semakin massif, perubahan perilaku konsumen yang semakin menyukai kemudahan dan kenyamanan dalam berbelanja [2]. Jumlah pengguna internet di Indonesia pada tahun 2023 mencapai 213 juta jiwa, jumlah ini meningkat sebesar 300% dibanding tahun 2013 [3]. Peningkatan jumlah pengguna internet dan mendorong pertumbuhan bisnis online [4]. Selain memberikan kemudahan, bisnis online juga menawarkan berbagai keuntungan dengan harga yang lebih kompetitif, pilihan yang lebih beragam, proses transaksi yang lebih mudah, serta pelayanan pelanggan yang lebih baik [5].

Bagi penjualan online makanan, perhitungan stok barang yang akurat sangat penting dilakukan untuk memenuhi permintaan pelanggan, kehabisan stok barang dapat menyebabkan kerugian bagi penjual [6]. Selain itu, perhitungan stok yang akurat penting untuk menjaga kualitas makanan, menghindari makanan menjadi basi atau rusak karena adanya kelebihan stok [7].

Dalam dunia data mining, hasil model asosiasi sering digunakan untuk interpretasi dalam menentukan pola pembelian pelanggan dan pola ini dapat digunakan untuk menyesuaikan stok barang [8]. Dengan menerapkan algoritma ini, pedagang retail dapat meningkatkan efisiensi dan efektivitas pengelolaan stok barang [9].

Pada data yang besar, algoritma asosiasi sering menampilkan nilai *lift* dan *confidence* yang rendah. Untuk itu perlu dilakukan penanganan data dengan melakukan pengurangan dimensi. Salah satu caranya adalah dengan melakukan *clustering* data. Terdapat penelitian yang menggunakan k-means, k-medoids dan fuzzy z-means dan selanjutnya diteruskan dengan FP-Growth. Hasilnya model yang menggunakan k-means menghasilkan aturan asosiasi dengan rata-rata *lift ratio* tertinggi, dan model yang menggunakan fuzzy c-means menghasilkan aturan asosiasi terbanyak [10].

Penelitian lainnya yang pernah dilakukan adalah dengan menggunakan k-medoids, dimana hasil penelitiannya menunjukkan itemset besar tanpa dilakukan *clustering* tidak ditemukan asosiasi pada nilai minimal *support* 5%, sedangkan setelah menggunakan k-medoids, ditemukan aturan asosiasi pada minimal *support* 5-9% [11].

Berdasarkan penelitian sebelumnya, maka dalam penelitian ini pada proses *clustering* akan digunakan algoritma lain, selain dari algoritma yang

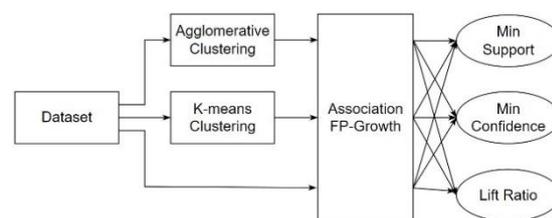
sudah pernah digunakan penelitian sebelumnya, yaitu *Agglomerative Clustering*. Tujuan dari penelitian ini adalah menemukan model yang tepat dalam penanganan data yang besar untuk mendapatkan aturan asosiasi dengan nilai minimal *support*, nilai *confidence* dan *lift ratio* yang tinggi dan valid, menggunakan salah satu algoritma *clustering* yang belum diuji pada penelitian sebelumnya yaitu *agglomerative clustering* untuk melihat bagaimana efeknya terhadap penanganan data yang besar apakah dapat meningkatkan nilai minimal *support*, nilai *lift* dan aturan asosiasi yang lebih banyak dibandingkan dengan *cluster* yang sudah pernah diteliti.

Hasil dari penelitian ini diharapkan dapat memberikan rekomendasi bagi penjual online yang memiliki variasi barang yang banyak dalam pengelolaan stok barang, serta strategi penjualan yang meningkatkan profit.

2. METODE PENELITIAN

Pada penelitian ini akan dilakukan tiga tahap penelitian pada *dataset*, yaitu dengan melakukan (1) penelitian dalam mencari asosiasi tanpa dilakukan *clustering* terlebih dahulu; (2) dilakukan *clustering* terlebih dahulu dengan menggunakan k-means dilanjutkan algoritma asosiasi menggunakan FP-Growth, dan (3) dilakukan *clustering* terlebih dahulu dengan menggunakan *agglomerative cluster*. Data yang di-*clustering* terlebih dahulu diharapkan dapat menangani variasi produk pada dataset. Untuk masing-masing algoritma, K-Means dan Agglomerative cluster, dipilih jumlah *cluster* terbaik berdasarkan nilai DBI. Selanjutnya, setelah dilakukan *clustering*, atas dataset dicari asosiasinya. Proses asosiasi dilakukan dengan dua tahap yaitu membuat *frequent itemsets* dan *association rules* dengan algoritma FP-Growth.

Tahapan penelitian dapat dilihat pada gambar dibawah ini.



Gambar 1. Tahapan penelitian

2.1. Tahapan Penelitian

Data set pada penelitian ini merupakan data transaksi penjualan makanan frozen dengan jumlah atribut sebanyak 5 kolom dan jumlah transaksi sebanyak 18.021 baris transaksi.

Tabel 1. Tampilan *Dataset*

Id_transaksi	tahun	nama_barang	jumlah_ruil	keterlibatan
PENJUALAN_1	2020	Egg Roll Chicken isi 3	1	1
PENJUALAN_1	2020	Cireng isi ayam pedas isi 10	1	1
PENJUALAN_1	2020	Spicy chicken isi 10	1	1
PENJUALAN_2	2020	Bakso cilok premium isi 20	3	1

Dataset dalam penelitian ini terdiri dari 202 jenis barang. Kolom keterlibatan merupakan jumlah transaksi yang mengandung barang tersebut. Jumlah riil merupakan jumlah total pembelian atas barang tersebut. Untuk lebih jelasnya dapat dilihat pada gambar dibawah ini.

Tabel 2. Jenis Barang

no	nama barang
1	Almond Milk Chocolate
2	Arem Mie isi 10
3	Arem Mie isi 5
4	Aren Latte
5	Ayam Ungkep
...	...
198	sisos solo isi 12
199	steak PO
200	teba
201	temulawak
202	udang kupas 250 gram

Selanjutnya dilakukan proses pemodelan pada *dataset* dengan hanya melakukan pemodelan algoritma FP-Growth tanpa menggunakan *clustering* dan pemodelan yang terlebih dahulu menggunakan *clustering*, baru setelahnya melakukan algoritma FP-Growth. Model tersebut dibandingkan *lift ratio*, minimal *support* dan minimal *confidencenya*.

2.2. Clustering Algorithm

K-Means *clustering* bekerja dengan cara membagi data poin ke dalam k klaster, di mana setiap data poin dikelompokkan ke dalam klaster dengan centroid terdekat. Centroid adalah titik pusat dari setiap klaster, yang dihitung sebagai rata-rata dari data poin yang dikelompokkan ke dalam klaster tersebut [12].

Agglomerative clustering adalah salah satu algoritma pengelompokan hierarkis (*Hierarchical Clustering*). Algoritma ini bekerja dengan menggabungkan dua *cluster* yang paling mirip pada setiap langkah, hingga semua data tercakup dalam satu *cluster*. Pada *agglomerative clustering*, setiap data awalnya dianggap sebagai satu *cluster*. Pada setiap langkah, dua *cluster* yang paling mirip akan digabungkan menjadi satu *cluster* baru. Proses ini diulangi hingga semua data tercakup dalam satu *cluster* [13].

Ada beberapa metode linkage yang dapat digunakan untuk menentukan dua *cluster* yang paling mirip pada *Hierarchical Clustering* [14]. Metode linkage yang paling umum digunakan adalah:

1. *Single linkage*: Dua *cluster* yang paling mirip adalah dua *cluster* yang memiliki jarak antar anggotanya paling kecil.
2. *Complete linkage*: Dua *cluster* yang paling mirip adalah dua *cluster* yang memiliki jarak antar anggotanya paling besar.
3. *Average linkage*: Dua *cluster* yang paling mirip adalah dua *cluster* yang memiliki rata-rata jarak antar anggotanya paling kecil.

4. *Ward linkage*: Dua *cluster* yang paling mirip adalah dua *cluster* yang memiliki varians antar anggotanya paling kecil.

Pilihan metode linkage yang digunakan akan mempengaruhi hasil pengelompokan. Misalnya, metode single linkage akan cenderung menghasilkan *cluster* yang lebih kecil, sedangkan metode complete linkage akan cenderung menghasilkan *cluster* yang lebih besar.

2.3. Proses Asosiasi

Asosiasi adalah salah satu algoritma data mining yang digunakan untuk menemukan hubungan antara *item-item* data dalam *dataset*. Algoritma ini sering digunakan untuk analisis data transaksional, seperti data penjualan, data kunjungan situs web, dan data penggunaan aplikasi. Asosiasi menggunakan dua metrik untuk mengukur pentingnya suatu hubungan antara *item-item* data yaitu *support*, *confidence*, dan *lift ratio*.

Support adalah presentasi kombinasi item tersebut dalam database. Rumus untuk menghitung nilai *support* dari dua item tersebut adalah:

$$Support(X) = |X| / |D| \quad (1)$$

Keterangan:

X adalah himpunan item

|X| adalah jumlah transaksi yang mengandung X

|D| adalah jumlah total transaksi

Nilai *support* menunjukkan seberapa sering X muncul dalam data. Nilai *support* yang tinggi menunjukkan bahwa X adalah kombinasi item yang umum terjadi. Dalam analisis asosiasi, *support* digunakan untuk menentukan apakah suatu aturan asosiasi valid atau tidak. Aturan asosiasi yang memiliki *support* di atas nilai minimum *support* yang ditentukan dianggap valid.

Nilai *confidence* menunjukkan seberapa sering Y muncul dalam transaksi yang mengandung X. Nilai *confidence* yang tinggi menunjukkan bahwa Y cenderung muncul setelah X. Kuat atau tidaknya hubungan antar-item dalam aturan asosiasi dapat diukur melalui *confidence*. Rumus untuk mendapatkan nilai *confidence* dari dua item yang dinilai memiliki hubungan tersebut adalah sebagai berikut:

$$Confidence(X \rightarrow Y) = support(X \cup Y) / support(X) \quad (2)$$

Keterangan:

X adalah himpunan item antecedent

Y adalah himpunan item consequent

|X U Y| adalah jumlah transaksi yang mengandung X dan Y

Lift ratio adalah salah satu parameter yang digunakan dalam analisis asosiasi untuk mengukur kekuatan suatu aturan asosiasi. *Lift ratio* menunjukkan seberapa besar peningkatan peluang

munculnya item *consequent* dalam transaksi yang mengandung item *antecedent*. Rumus *lift ratio* adalah sebagai berikut:

$$\text{Lift}(X \rightarrow Y) = \frac{\text{confidence}(X \rightarrow Y)}{\text{confidence}(Y)} \quad (3)$$

Keterangan:

X adalah himpunan item antecedent

Y adalah himpunan item consequent

confidence(X → Y) adalah nilai *confidence* aturan asosiasi X → Y

confidence(Y) adalah nilai *confidence* aturan asosiasi Y

Frequent Pattern Growth (FP-Growth) adalah algoritma data mining yang digunakan untuk menemukan pola asosiasi dalam data transaksional. Algoritma ini merupakan pengembangan dari algoritma Apriori [15]. Proses growth digunakan untuk menemukan frequent itemset dari FP-Tree FP-Growth merupakan algoritma yang efisien untuk menemukan pola asosiasi dalam data transaksional. Algoritma ini memiliki beberapa kelebihan, seperti efisien dalam penggunaan memori dan waktu serta dapat digunakan untuk menemukan pola asosiasi dengan ukuran yang besar. Dengan menggunakan FP-Tree, algoritma FP-growth dapat langsung mengekstrak frequent Itemset dari FP-Tree [16].

3. HASIL DAN PEMBAHASAN

3.1. Clustering

Data awal terdiri dari 18.021 baris transaksi dan 5 kolom atribut. Kolom atribut terdiri dari *id_transaksi*, *tahun*, *nama_barang*, *jumlah_riil* dan *keterlibatan*. Sebelum dilakukan *clustering*, kolom *keterlibatan* dan *jumlah_riil* dilakukan agregasi berdasarkan *nama_barang*. Hasilnya menjadi sebuah tabel dengan 202 baris dan 3 kolom sebagaimana ditampilkan pada Tabel 3.

Tabel 3. Hasil Agregasi Berdasarkan Nama Barang

no	nama_barang	keterlibatan	jumlah_riil
1	Almond Milk Chocolate	50	201
2	Arem Mie isi 10	5	5
3	Arem Mie isi 5	5	5
4	Aren Latte	1	1
5	Ayam Ungkep	8	8
...
198	sisis solo isi 12	23	26
199	steak PO	1	14
200	teba	69	102
201	temulawak	3	3
202	udang kupas 250 gram	4	4

Data hasil agregasi kemudian dihitung nilai tengahnya (Median) dan didapatkan median untuk *keterlibatan* adalah 30 dan median untuk *jumlah_riil* adalah 42. Nilai median *keterlibatan* 30 dan *jumlah*

riil 42 merupakan nilai yang sangat kecil dibandingkan total keseluruhan transaksi sebanyak 18.021 transaksi.

Setelah dianalisa secara mendalam, ternyata penjualan dengan keterlibatan kurang dari sama dengan 30 merupakan penjualan barang yang tidak rutin. Barang tersebut merupakan barang yang dijual berdasarkan permintaan khusus dari pembeli, sehingga penjualan barang tersebut perlu dikeluarkan dari data yang akan dilakukan proses *clustering*

Setelah dikurangi produk yang dijual secara tidak rutin maka menghasilkan 95 jenis barang sebagaimana ditampilkan pada Tabel 4. Dari proses agregasi dapat dilihat bahwa Dimsum Box isi 25 merupakan produk dengan penjualan terbanyak yaitu sebanyak 5.954 dan terendah adalah Bakso Aci Mozarella sebanyak 60.

Tabel 4. Hasil Agregasi penjualan barang rutin

no	nama_barang	keterlibatan	jumlah_riil
1	Dimsum BOX isi 25	2449	5954
2	Beef slice non lemaks 500gram	624	1114
3	Egg Roll Chicken isi 3	446	705
4	Dimsum MENTAI isi 6	428	738
5	Tahu bakso premium isi 10	405	639
...
91	Dimsum matang isi 4 pcs	39	92
92	dimsum matang isi 10	33	47
93	Saikoro 250 gram	33	77
94	Canai ori ?isi 5 lembar	32	48
95	Bakso Aci Mozarella	32	60

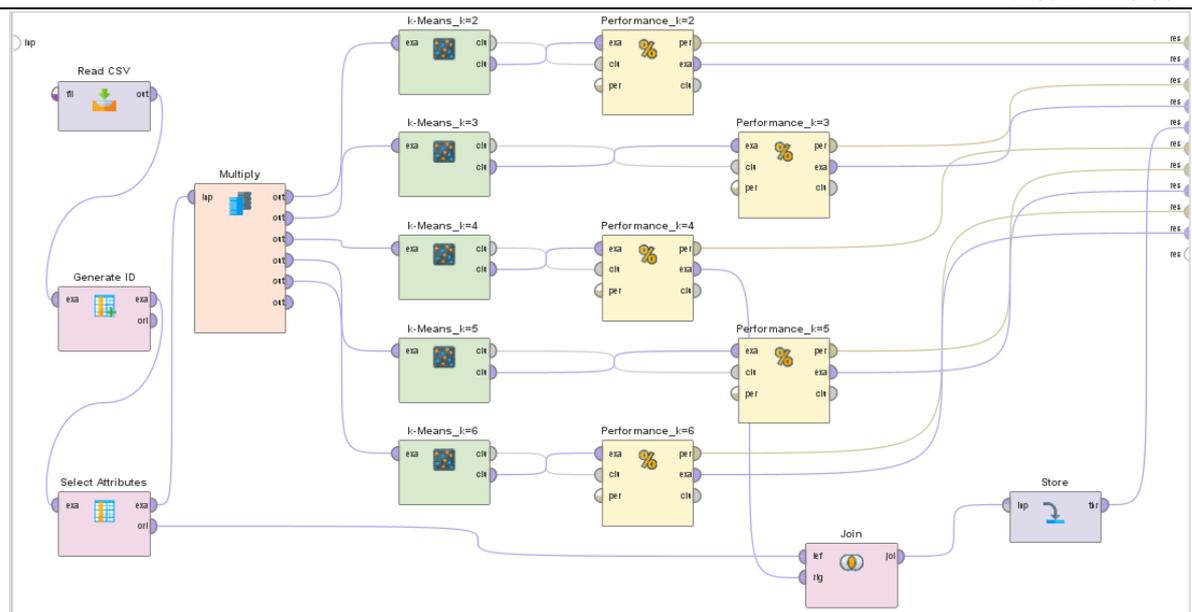
Tahap selanjutnya adalah *clustering* menggunakan algoritma K-Means dengan bantuan Aplikasi RapidMiner. Proses *clustering* dilakukan beberapa kali dengan jumlah *cluster* 2, 3, 4, 5, dan 6.

Masing-masing *cluster* yang terbentuk kemudian dilakukan evaluasi performanya dengan menggunakan nilai rata-rata jarak centroid dan nilai Davies Boulding Index (DBI). Model *Clustering* menggunakan algoritma K-Means ditampilkan pada Gambar 2. *Cluster* terbaik adalah *cluster* dengan nilai DBI terkecil yang menunjukkan bahwa *cluster* memiliki karakteristik yang baik dalam hal kesamaan antar anggotanya dan perbedaan antara *cluster* yang optimal.

Hasil evaluasi model *clustering* dengan menggunakan algoritma K-Means menampilkan hasil DBI terendah pada jumlah *cluster* k=4 dengan nilai DBI 0,368 dan nilai rata-rata jarak centroid sebesar 3189,097. Hasil evaluasi secara lengkap untuk algoritma K-Means ditampilkan pada Tabel 5.

Tabel 5. Evaluasi model K-means

Jumlah k	Nilai Rata-rata Jarak Centroid	DBI
2	14.228,458	0,555
3	6.204,190	0,513
4	3.189,097	0,368
5	1.548,027	0,370
6	993,009	0,405



Gambar 2. Clustering K-Means

Model *clustering* yang kedua menggunakan algoritma *Agglomerative Clustering* dan diterapkan pada jumlah *cluster* k 2, 3, 4, 5 dan 6. Didapati hasil terbaik adalah k=3. Dimana pada jumlah *cluster* 3, tidak terdapat *cluster* yang hanya berisi satu item saja. Jumlah anggota masing-masing *cluster* untuk k=3 adalah *cluster_0* terdiri dari 88 items, *cluster_1* terdiri dari 2 items dan *cluster_2* terdiri dari 5 items.

3.2. Asosiasi

Tahapan kedua setelah dilakukan *clustering* adalah melakukan asosiasi. Proses asosiasi dilakukan dengan dua tahap yaitu membuat *frequent itemsets* dan dari *frequent itemsets* tersebut dibuat *association*

rules dengan algoritma FP-Growth. Proses asosiasi menggunakan tiga *dataset* yaitu *dataset* yang tidak dilakukan *clustering*, *dataset* hasil *cluster* K-Means dan *dataset* hasil *cluster* *Agglomerative Clustering*. Untuk masing-masing *dataset* juga ditentukan nilai *support* minimalnya 0,01 dan nilai *confidencenya* 0,1.

Hasil pemodelan asosiasi yang pertama menggunakan *dataset* yang tidak dilakukan *clustering* terlebih dahulu menghasilkan 10 *frequent itemsets*. Semua *itemsets* yang dihasilkan terdiri dari 1 item, dengan nilai *support* terendah dengan 0,01 dan nilai *confidence* terendah 0,206. Sedangkan nilai *support* tertinggi adalah 0,017 dan nilai *confidencenya* 0,414.

Tabel 6. Asosiasi Dataset Tanpa Clustering

PREMISES	CONCLUSION	SUPPORT	CONFIDENCE	LIFT
Nugget ayam 1 kg	Dimsum BOX isi 25	0,010	0,206	0,659
Bakso sapi isi 50	Dimsum BOX isi 25	0,011	0,237	0,758
Egg Roll Chicken isi 3	Shrim roll isi 10	0,014	0,249	7,273
Tahu bakso premium isi 10	Dimsum BOX isi 25	0,013	0,249	0,798
Bakso cilok premium isi 20	Dimsum BOX isi 25	0,012	0,259	0,828
Rujak cireng isi 20	Dimsum BOX isi 25	0,012	0,261	0,833
Egg Roll Chicken isi 3	Dimsum BOX isi 25	0,017	0,298	0,954
Singkong keju D9	Dimsum BOX isi 25	0,010	0,352	1,126
Beef slice lemak 500 gram	Beef slice non lemak 500gram	0,015	0,389	4,879
Shrim roll isi 10	Egg Roll Chicken isi 3	0,014	0,414	7,273

Dari Tabel 6 dapat dilihat bahwa nilai *lift* yang lebih dari 1 sebanyak 4 aturan asosiasi. Artinya, dengan *confidence* lebih dari 0,2, hanya 4 aturan asosiasi yang memiliki kecenderungan untuk benar dibandingkan dengan kejadian acak.

Hasil pemodelan asosiasi yang kedua dapat menggunakan *dataset* hasil *cluster* K-Means menghasilkan 18 *frequent itemsets* yang memiliki *support* lebih dari 0,01 dan nilai *confidence* lebih dari 0,1. Nilai *support* tertinggi adalah 0,036 dan nilai

confidence tertinggi adalah 0,222. Hasil pemodelan asosiasi yang kedua ini mendapatkan 11 aturan dengan nilai *lift* lebih dari 1. Secara lengkap hasil asosiasi untuk *dataset cluster* K-Means ditampilkan pada Tabel 7, dimana kolom C merupakan kolom *Cluster*, kolom PREM merupakan kolom *Premises*, kolom CONC merupakan kolom *Conclusion*, kolom SUP merupakan kolom *Support*, kolom CONF merupakan kolom *Confidence* dan kolom LIFT.

Tabel 7. Asosiasi *Dataset Cluster K-Means*

C	PREM	CONC	SUP	CONF	LIFT
0	no association rule				
1	Ikan dori fillet 1 kg	Eby furai isi 12	0,011	0,117	1,673
1	Kentang goreng 1 kg	Bakso sapi isi 50	0,011	0,119	1,182
1	Nugget ayam 1 kg	Sosis breakfast bernardi 1 kg	0,014	0,130	1,940
1	Shrim roll isi 10	Eby furai isi 12	0,011	0,146	2,089
1	Eby furai isi 12	Ikan dori fillet 1 kg	0,011	0,151	1,673
1	Eby furai isi 12	Shrim roll isi 10	0,011	0,155	2,089
1	Sosis breakfast bernardi 1 kg	Nugget ayam 1 kg	0,014	0,212	1,940
1	Shrim roll isi 10	Kaki naga isi 10	0,016	0,213	2,982
1	Kaki naga isi 10	Shrim roll isi 10	0,016	0,222	2,982
2	Beef slice non lemak 500gram	Dimsum BOX isi 25	0,036	0,171	0,208
3	Bakso cilok premium isi 20	Rujak cireng isi 20	0,021	0,120	0,708
3	Piscok lumer isi 10	Tahu bakso premium isi 10	0,019	0,122	0,628
3	Rujak cireng isi 20	Bakso cilok premium isi 20	0,021	0,125	0,708
3	Pempek Palembang isi 9	Bakso cilok premium isi 20	0,014	0,125	0,710
3	Tahu bakso premium isi 10	Bakso cilok premium isi 20	0,025	0,128	0,730
3	Bakso cilok premium isi 20	Tahu bakso premium isi 10	0,025	0,142	0,730
3	Rujak cireng isi 20	Piscok lumer isi 10	0,033	0,195	1,243
3	Piscok lumer isi 10	Rujak cireng isi 20	0,033	0,210	1,243

Hasil pemodelan asosiasi dari *dataset cluster Agglomerative Clustering* menghasilkan 6 frequent itemsets dengan nilai *support* diatas 0,01 dan nilai *confidence* diatas 0,1. Dimana nilai tertinggi *support* adalah 0,036 dan nilai tertinggi *confidence* adalah

0,21. Untuk pemodelan asosiasi *dataset* ini hanya menghasilkan 1 aturan yang memiliki nilai *lift* diatas 1. Hasil pemodelan asosiasinya ditampilkan pada Tabel 8, dimana keterangan untuk masing-masing kolom sama dengan Tabel 7.

Tabel 8. Asosiasi *Dataset Cluster Agglomerative*

C	PREM	CONC	SUP	CONF	LIFT
0	Piscok lumer isi 10	Rujak cireng isi 20	0,012	0,210	3,490
1	Beef slice non lemak 500gram	Dimsum BOX isi 25	0,036	0,171	0,208
2	Tahu bakso premium isi 10	Egg Roll Chicken isi 3	0,026	0,106	0,394
2	Pempek Palembang isi 9	Bakso cilok premium isi 20	0,018	0,125	0,564
2	Tahu bakso premium isi 10	Bakso cilok premium isi 20	0,031	0,128	0,579
2	Bakso cilok premium isi 20	Tahu bakso premium isi 10	0,031	0,142	0,579

4. DISKUSI

Bagian ini menjelaskan hasil yang didapatkan dari seluruh tahapan penelitian. Untuk memberikan

kemudahan dalam memilih aturan asosiasi terbaik, dibuat rekap dalam bentuk tabel sebagaimana ditampilkan pada Tabel 9.

Tabel 9. Rekap Perbandingan Asosiasi

<i>Dataset</i>	Jumlah <i>Cluster</i>	Sup	Conf	Jumlah Aturan Asosiasi	Nilai <i>Lift</i> Diatas 1
Tanpa <i>Cluster</i>	-	0,01	0,1	10	4
<i>Cluster K-Means</i>	4			18	11
<i>Cluster Agglomerative Clustering</i>	3			6	1

Model asosiasi dengan menggunakan algoritma FP-Growth telah diterapkan terhadap tiga *dataset*, yaitu *dataset* tanpa *cluster*, *dataset cluster K-Means* dan *dataset cluster Agglomerative*. Jumlah *cluster* yang dihasilkan untuk masing-masing model dengan evaluasi terbaik juga berbeda dimana *cluster* terbaik untuk model K-Means adalah k=4, sedangkan *cluster* terbaik untuk model Agglomerative Clustering adalah k=3. Minimum *support* yang digunakan untuk penyusunan frequent itemsets sama untuk ketiga *dataset* yaitu sebesar 0,01, begitu pula untuk minimum *confidence* juga sama yaitu 0,1. Namun jumlah aturan asosiasi beserta nilai *lift* untuk ketiga *dataset* tersebut berbeda-beda. Jumlah aturan asosiasi terbanyak dihasilkan dari *dataset cluster K-Means* yaitu sebanyak 18 aturan asosiasi, kemudian *dataset* tanpa *cluster* sebanyak 10 aturan asosiasi dan *dataset cluster Agglomerative Clustering* dengan 6 asosiasi. Hal tersebut juga sebanding dengan jumlah aturan

asosiasi dengan nilai *lift* lebih dari 1, dimana *cluster K-Means* terbanyak yaitu 11 aturan asosiasi dengan nilai *lift* lebih dari 1, dapat dilihat pada tabel 7. Sehingga dapat disimpulkan bahwa asosiasi terbaik didapatkan dari *dataset cluster K-Means*.

Bila dibandingkan dengan penelitian sebelumnya yang mencari optimalisasi asosiasi dengan melakukan algoritma *clustering* pada *dataset* menggunakan K-Means, K-Medoids dan Fuzzy c-means, proses optimisasi rekomendasi produk pada penelitian ini menggunakan algoritma *clustering* yang berbeda yaitu algoritma Agglomerative Clustering dan algoritma K-Means, baru hasilnya dicari aturan asosiasinya dengan menggunakan algoritma FP-Growth. Algoritma Agglomerative clustering menggunakan teknik *clustering* yang berbeda dengan algoritma K-Means. Dimana pada algoritma Agglomerative Clustering, pembentukan *cluster* dengan mengelompokkan data menjadi

sebuah hierarki, sedangkan pada algoritma K-Means pembentukan *cluster* dengan menghitung jarak ke titik pusat *cluster*.

Dari model terbaik, didapatkan 11 aturan asosiasi dengan nilai *lift* diatas 1. Penjual online dapat memfokuskan pada pasangan item dari aturan asosiasi tersebut untuk meningkatkan penjualan, misalnya dengan menjual pasangan produk dengan promo tertentu. Aturan asosiasi secara lengkap ditampilkan pada Gambar 3.

IF [Ikan dori fillet 1 kg], THEN [Eby furai isi 12]
IF [Kentang goreng 1 kg], THEN [Bakso sapi isi 50]
IF [Nuget ayam 1 kg], THEN [Sosis breakfast bernardi 1 kg]
IF [Shrim roll isi 10], THEN [Eby furai isi 12]
IF [Eby furai isi 12], THEN [Ikan dori fillet 1 kg]
IF [Eby furai isi 12], THEN [Shrim roll isi 10]
IF [Sosis breakfast bernardi 1 kg], THEN [Nuget ayam 1 kg]
IF [Shrim roll isi 10], THEN [Kaki naga isi 10]
IF [Kaki naga isi 10], THEN [Shrim roll isi 10]
IF [Rujak cireng isi 20], THEN [Piscok lumer isi 10]
IF [Piscok lumer isi 10], THEN [Rujak cireng isi 20]

Gambar 3. Aturan Asosiasi

5. KESIMPULAN

Dari hasil penelitian yang sudah dilakukan, diketahui pada *dataset cluster* K-Means dan *dataset cluster* Agglomerative menghasilkan nilai *support* lebih tinggi dibandingkan *dataset* tanpa *cluster* yaitu sebesar 0,03. Kemudian, jumlah aturan asosiasi terbanyak dihasilkan dari *dataset cluster* K-Means yaitu sebanyak 18 aturan asosiasi, dengan 11 aturan asosiasi yang nilai *lift* lebih dari 1. Dapat disimpulkan *dataset cluster* K-Means memberikan aturan asosiasi yang dapat menjadi rekomendasi terbaik bagi penjual online untuk melakukan pengelolaan stok barang, serta berbagai rencana strategi penjualan dengan tujuan peningkatan profit.

Penelitian ini telah memberikan informasi tentang kombinasi metode *clustering* dan asosiasi dalam penentuan rekomendasi produk, namun perlu dilakukan penelitian dengan menggunakan metode lain untuk lebih meningkatkan nilai *support*, *confidence* dan *lift* aturan asosiasi produk dengan jenis produk yang sangat beragam.

DAFTAR PUSTAKA

- [1] A. Supriyanto and K. F. Hana, "Strategi Pengembangan Desa Digital Untuk Meningkatkan Produktivitas UMKM," *BISNIS: Jurnal Bisnis dan Manajemen Islam*, vol. 8, no. 2, p. 199, Dec. 2020, doi: 10.21043/bisnis.v8i2.8640.
- [2] P. Iffa Rosada, M. Syifa Amin Widigdo, E. Syariah, F. Agama Islam, and U. Muhammadiyah Yogyakarta, "Analisis Perilaku Konsumen Muslim Dalam Belanja Melalui Shopee Saat Pandemi Covid-19 Berdasarkan Perspektif Islam," *Jurnal Ilmiah Ekonomi Islam*, vol. 9(02), 2023, doi: 10.29040/jiei.v9i2.7123.
- [3] F. L. Dasri and E. Suwarni, "Analisis Disrupsi Media Digital Terhadap Industri Televisi Indonesia," 2023.
- [4] L. Marlinah, "Mendorong Pertumbuhan Ekonomi Indonesia Melalui Penguatan Sektor Ekonomi Digitalpreneur Dan Creativepreneur," 2019.
- [5] M. S. Hotana, "Industri E-Commerce Dalam Menciptakan Pasar Yang Kompetitif Berdasarkan Hukum Persaingan Usaha," *Jurnal Hukum Bisnis Bonum Commune*, vol. 1, no. 1, 2018.
- [6] L. Sari Marita and I. Darwati, "Prediksi Persediaan Barang Menggunakan Metode Weighted Moving Average, Exponential Smoothing dan Simple Moving Average," *Jurnal TEKNO KOMPAK*, vol. 16, no. 1, 2022.
- [7] R. P. Megatama, J. M. Maligan, and R. Septifani, "Studi Inventory Control Pada National Fullfillment Center Cikarang Di PT XYZ Indonesia," *PROSIDING SEMINAR NASIONAL INSTIPER*, vol. 1, no. 1, pp. 314–322, Jul. 2022, doi: 10.55180/pro.v1i1.268.
- [8] A. Maulana and A. A. Fajrin, "Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma FP-Growth Pada Data Transaksi Penjualan Spare Part Motor," *Kumpulan jurnal Ilmu Komputer*, vol. 5, no. 1, 2018.
- [9] I. Nurrohmat, O. Nurdiawan, and A. Bahtiar, "Implementasi Algoritma FP-Growth Untuk Menunjang Keputusan Persediaan Barang Di CV Indotech Jaya Sentosa Kota Cirebon," *Jurnal Sistem Informasi dan Manajemen*, vol. 10, no. 2, 2022, [Online]. Available: <https://ejournal.stmikgici.ac.id/>
- [10] F. Nuraeni, D. Tresnawati, Y. Handoko Agustin, and G. Fauzi, "Optimization of Market Basket Analysis Using Centroid-Based Clustering Algorithm And FP-Growth Algorithm," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 6, pp. 1581–1590, Dec. 2022, doi: 10.20884/1.jutif.2022.3.6.399.
- [11] I. Syukra, A. Hidayat, and M. Z. Fauzi, "Implementation of K-Medoids and FP-Growth Algorithms for Grouping and Product Offering Recommendations," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 2, no. 2, p. 107, Nov. 2019, doi: 10.24014/ijaidm.v2i2.8326.
- [12] S. M. Dewi, A. P. Windarto, I. S. Damanik, and H. Satria, *Analisa Metode K-Means pada Pengelompokan Kriminalitas Menurut Wilayah*. Seminar Nasional Sains &

Teknologi Informasi (SENSASI), 2019.
[Online]. Available:
[http://prosiding.seminar-
id.com/index.php/sensasi/issue/archivePage|
620](http://prosiding.seminar-id.com/index.php/sensasi/issue/archivePage|620)

- [13] E. Widodo, P. Ermayani, L. N. Laila, and A. T. Madani, "Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Analisis Hierarchical Agglomerative Clustering (Indonesian Province Grouping Based on Poverty Level Using Hierarchical Agglomerative Clustering Analysis)," 2021.
- [14] Vijaya, S. Sharma, and N. Batra, *Comparative Study of Single Linkage, Complete Linkage, and Ward Method of Agglomerative Clustering*. 2019.
- [15] F. Nuraeni, D. Tresnawati, Y. Handoko Agustin, and G. Fauzi, "Optimization of Market Basket Analysis Using Centroid-Based Clustering Algorithm and FP-Growth Algorithm," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 6, pp. 1581–1590, Dec. 2022, doi: 10.20884/1.jutif.2022.3.6.399.
- [16] B. Septia Pranata and D. Putro Utomo, "Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor (Study Kasus Bengkel Sinar Service)," *Bulletin of Information Technology (BIT)*, vol. 1, no. 2, pp. 83–91, 2020.