

Implementation of K-Medoids and FP-Growth Algorithms for Grouping and Product Offering Recommendations

¹Imaduddin Syukra, ²Assad Hidayat, ³Muhammad Zakiy Fauzi

^{1,2,3}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi

^{1,2,3}Puzzle Research Data Technology (Predatech), Fakultas Sains dan Teknologi

Universitas Islam Negeri Sultan Syarif Kasim Riau

Email: ¹imaduddinsyukra@gmail.com, ²hidayatassad@gmail.com, ³zakiyfauzi@gmail.com

Article Info

Article history:

Received Sep 06th, 2019

Revised Nov 11th, 2019

Accepted Nov 21th, 2019

Keyword:

Association Rule

Clustering

Data Mining

FP-Growth

Heart Attack

ABSTRACT

212 Mart Rambutan Street on Pekanbaru City is a company engaged in retail. Meeting the needs of consumers and making the right decision in determining the sales strategy is a must. One way to find out market conditions is to observe sales transaction data using data mining. The data mining method commonly used to analyze market basket (Market Basket Analysis) is the Association Rule. The Association Rule can provide product recommendations and promotions, so that the marketing strategy is more targeted and the items promoted are the customer's needs. At 212 Mart, the determination of product promotion is obtained from the analysis of sales transaction data reports, which are based on the most sold products and the expiration date. Often the product being promoted does not fit the customer's needs. The purpose of this study is to apply the K-Medoids algorithm for clustering on FP-Growth in producing product recommendation rules on a large number of datasets so that they can provide technical recommendations / new ways to the 212 Mart in determining product promotions. The results obtained are from the experiments the number of clusters 3 to 9 obtained optimal clusters of 3 clusters based on the validity test of the Davies Bouldin Index with a value of 0.678. With a minimum support value of 5% - 9% and a minimum value of 50% confidence, the result is that the Association Rule is found only in cluster 3 with 5 rules.

Copyright © 2019 Puzzle Research Data Technology

Corresponding Author:

Muhammad Zakiy Fauzi,

Program Studi Sistem Informasi,

Universitas Islam Negeri Sultan Syarif Kasim Riau,

Jl. HR. Soebrantas No. 155 Simpang Baru, Tampan Pekanbaru, Riau - Indonesia, 28293

Email: zakiyfauzi@gmail.com

DOI: <http://dx.doi.org/10.24014/ijaidm.v2i2.8326>

1. PENDAHULUAN

Persaingan bisnis dalam perdagangan dunia melalui ekonomi pasar bebas yang diikuti dengan kemajuan teknologi informasi membawa perusahaan pada tingkat persaingan yang semakin ketat dan semakin terbuka dalam memenuhi tuntutan pelanggan yang juga semakin tinggi. Sehingga memaksa para pelakunya untuk senantiasa memikirkan strategi-strategi dan terobosan bagaimana untuk tetap bertahan dan mengembangkan pangsa pasar bisnis mereka.

212 Mart Jalan Rambutan Kota Pekanbaru merupakan salah satu perusahaan yang bergerak dibidang retail. Memenuhi kebutuhan konsumen setiap harinya dan mengambil keputusan yang tepat dalam menentukan strategi penjualan merupakan suatu keharusan. Untuk itu, dibutuhkan sumber informasi yang cukup banyak untuk dapat dianalisis lebih lanjut. Pada bisnis retail, salah satu cara untuk mengetahui bagaimana kondisi pasar (pelanggan) saat ini adalah melakukan pengamatan data transaksi penjualan. Data transaksi penjualan ini disimpan dalam basis data server yang kemudian diolah sehingga menghasilkan sebuah informasi berupa laporan penjualan dan laporan laba rugi. Namun, pengolahan data lebih lanjut sangat dibutuhkan agar dapat menghasilkan informasi baru [8]. Pengolahan data yang lebih lanjut, dapat dilakukan dengan proses ekstraksi dan identifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari data dengan jumlah yang besar

(database) dengan menggunakan teknik-teknik tertentu seperti teknik statistic, matematika, kecerdasan buatan, dan machine learning yang biasa disebut dengan Data mining [14]. Banyak metode yang terdapat pada data mining, salah satunya yaitu Association Rule. Association Rule biasa digunakan untuk menganalisis keranjang pasar (Market Basket Analysis) [6]. Analisis keranjang pasar (Market Basket Analysis) dapat membantu para pelaku usaha retail dalam memberikan dukungan keputusan ilmiah dengan melakukan Association Rule Mining antara barang-barang yang dibeli secara bersamaan oleh pelanggan [25]. Hal ini dapat membantu memberikan rekomendasi dan promosi produk, sehingga menghasilkan strategi pemasaran yang lebih tepat sasaran dan menarik lebih banyak pelanggan karena barang yang dipromosikan merupakan barang yang dibutuhkan oleh pelanggan. Penentuan pola pembelian yang kurang akurat dapat mengakibatkan kebijakan rekomendasi dan promosi produk menjadi tidak tepat sasaran [23].

Pada perusahaan 212 Mart Jalan Rambutan Kota Pekanbaru, penentuan promosi produk didapatkan dari analisis laporan data transaksi penjualan, yaitu berdasarkan pada produk yang paling laku. Data transaksi penjualan ini, dievaluasi 1 kali dalam 3 bulan. Sehingga, hasil yang didapatkan tidak efisien. Karena, untuk produk yang paling laku, sering kali promo yang ditawarkan tidak sesuai dengan kondisi. Contohnya, produk yang paling laku ditanggal tanggal tertentu seperti pada awal bulan, ditawarkan di akhir bulan, begitu juga sebaliknya. Selain itu, penentuan promosi produk juga didapat berdasarkan tanggal kedaluwarsa produk. Produk yang memiliki sisa waktu dalam kurun waktu 3 bulan menjelang kedaluwarsa, akan diadakan promo potongan harga. Sering kali produk-produk yang dipromosikan justru tidak sesuai dengan kebutuhan pelanggan.

Penelitian Terbaru dilakukan oleh Mustakim dan kawan-kawan (2018) dengan menerapkan teknik asosiasi yaitu dengan algoritma Apriori dan FP-Growth. Dari kedua algoritma tersebut, terdapat dua hal yang ingin diketahui dari analisis dataset penjualan pada Berkah Mart, yaitu proses rule yang akan dihasilkan dari masing-masing algoritma dan waktu yang dicapai masing-masing algoritma dalam membentuk rule. Hasil yang didapatkan adalah FP-Growth lebih efektif dalam menemukan frequent itemset dan menemukan pola pembelian pelanggan dibanding dengan Apriori [15].

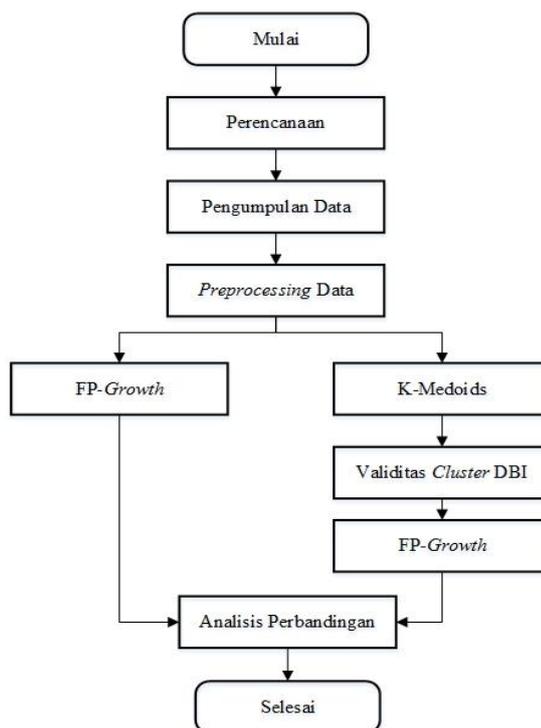
Dalam menganalisis keranjang pasar menggunakan Association Rules, ditemukan sejumlah masalah teknis umum yang berhubungan dengan teknik rekomendasi. Itemset besar cenderung diabaikan oleh aturan asosiasi, dan kurang tepatnya rekomendasi item dikarenakan tidak tersedianya informasi tentang produk retail [6], sehingga hasil yang diperoleh menjadi kurang akurat untuk data yang besar. Untuk itu, dilakukan clustering terhadap atribut-atribut yang ada guna membentuk kelompok atribut yang sama, setelah itu menentukan pola asosiasi pada masing-masing kelompok yang sudah terbentuk [18], sehingga mempermudah dalam proses pencarian rekomendasi produk.

Clustering adalah teknik yang efektif untuk analisis data eksploratori, dan telah diterapkan di berbagai bidang. Secara umum, metode untuk melakukan clustering dapat dikategorikan menjadi empat, yaitu: partitioning, hierarchical, grid-based and model-based. Contoh dari metode partitioning adalah K-Means dan K-Medoids [24]. Kelemahan K-Means terdapat pada penentuan centroid awal pada cluster. Hasil cluster yang terbentuk tergantung pada inisiasi nilai centroid awal cluster yang diberikan. Sehingga menyebabkan hasil cluster berupa solusi yang bersifat local optimal [19]. Selain itu, algoritma ini sensitif terhadap outlier karena objek dengan nilai yang sangat besar secara substansial dapat mendistorsi distribusi data. Untuk itu, dalam mengambil nilai rata-rata dari objek pada sebuah cluster sebagai titik acuan, medoid dapat diandalkan, yang mana merupakan objek yang paling terpusat dalam sebuah cluster. Menemukan k cluster dalam n objek dengan pertama kali secara arbitrarilly menemukan wakil dari objek (medoid) untuk tiap-tiap cluster, merupakan strategi dasar dari algoritma clustering K-Medoids [24].

Berdasarkan pembahasan penelitian sebelumnya diatas, pada penelitian ini akan menggunakan algoritma K-Medoids untuk clustering pada data penjualan dan menerapkan algoritma FP-Growth untuk pendekatan asosiasi pada setiap cluster yang telah dibentuk. Sehingga diharapkan untuk dapat memberikan rekomendasi produk yang lebih akurat kepada pelanggan dikarenakan dataset yang akan diasosiasi menjadi lebih kecil.

2. METODOLOGI PENELITIAN

Berikut ini akan membahas tentang metodologi pada penelitian yang akan dilakukan. Adapun langkah-langkah yang ditempuh dalam penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

Tahap perencanaan dilakukan agar tujuan penelitian lebih jelas dan terarah, yang terdiri dari Identifikasi Masalah, Perumusan Masalah, Menentukan Tujuan, Menentukan Batasan Masalah dan Studi Pustaka. Kemudian dilakukan studi literatur dan observasi yang merupakan bagian dari tahap pengumpulan data. Pada tahap Preprocessing dimulai dengan seleksi data, cleaning, transformasi dan normalisasi data. Analisa dan hasil terdiri dari analisis pendahuluan, analisis cluster dan asosiasi menggunakan metode K-Medoids dan FP-Growth, dan membandingkan rule terbaik yang terbentuk antara sebelum dilakukannya cluster dan sesudah dilakukannya cluster dalam penerapan algoritma FP-Growth.

2.1. Data Mining

Data mining adalah proses untuk menemukan informasi yang bermanfaat yang tersembunyi dalam data dengan jumlah yang besar, serta membantu dalam pengambilan keputusan bagi stakeholder dalam membuat keputusan, guna mengembangkan bisnisnya. Teknologi data mining menggabungkan kecerdasan buatan (Artificial Intelligence), mesin pembelajaran (Machine Learning), pengenalan pola (Pattern Recognition), teknik visualisasi, statistik [13] dan sistem manajemen basis data [7].

2.2. Analisis Keranjang Pasar

Analisis Keranjang Pasar (Market Basket Analysis), atau juga dikenal sebagai association rule mining atau analisis afinitas, adalah salah satu teknik dalam data mining [6] yang digunakan untuk mengidentifikasi hubungan antara kelompok produk, item, atau kategori. Analisis keranjang belanja awalnya berasal dari bidang pemasaran. Analisis Keranjang Pasar memudahkan peneliti untuk memahami perilaku pelanggan melalui produk mana yang dibeli secara bersamaan [1].

2.3. K-Medoids

Clustering adalah suatu proses pengelompokan kumpulan data ke dalam kelompok-kelompok sehingga objek-objek yang terdapat dalam suatu kelompok memiliki kesamaan yang tinggi dan berbeda dengan objek yang berada dalam kelompok lain [10]. Clustering bertujuan untuk mempartisi dataset yang diberikan ke dalam beberapa cluster yang mana dataset yang berada dalam suatu cluster memiliki kemiripan dan berbeda dari dataset yang berada di cluster lain [20] [21]. Pada algoritma K-Medoids objek yang menjadi perwakilan yang disebut medoid dianggap sebagai pengganti centroid (pusat dari cluster). Karena menggunakan objek yang paling terpusat dalam sebuah cluster, K-Medoids dianggap kurang sensitif terhadap outlier jika dibandingkan dengan algoritma K-means [16].

2.4. Davies Bouldien Index (DBI)

Konsep dasar dari DBI adalah untuk meminimalkan jarak rata-rata antar cluster yang terdekat [3]. Tujuan dari DBI adalah untuk mengidentifikasi kumpulan kelompok yang mirip tetapi terpisah jauh [5].

2.5. Association Rules

Asosiasi (Association Rules) adalah teknik data mining untuk menemukan aturan asosiatif antara kombinasi item. Untuk mengetahui kepentingan suatu aturan asosiasi dapat diketahui dengan dua parameter, yaitu support dan confidence. Support yaitu persentase kombinasi item tersebut dalam database. Sedangkan confidence yaitu kuatnya hubungan antara item dalam association rule [4].

2.6. FP-Growth

FP-Growth merupakan pengembangan dari algoritma Apriori. Pada algoritma FP-Growth, kekurangan dari algoritma Apriori diperbaiki dengan menghilangkan candidate generation, karena dalam FP-Growth digunakan konsep pembangunan pohon, yang biasa disebut FP-Tree dalam pencarian frequent itemset. Selain itu, scanning database hanya dilakukan sebanyak dua kali. Sehingga membuat algoritma FP-Growth memiliki performa yang lebih cepat jika dibandingkan dengan algoritma Apriori [9] [22] [25].

3. HASIL DAN ANALISIS

3.1. Analisis Pendahuluan

Pada tahapan analisa ini akan membahas mengenai data transaksi penjualan 212 Mart Jalan Rambutan Kota Pekanbaru pada 21 Oktober 2018 sampai dengan 12 April 2019 dan proses clustering menggunakan metode K-Medoids untuk pengelompokan produk dan proses asosiasi menggunakan metode FP-Growth untuk menemukan pola pembelian pelanggan metode FP-Growth untuk menemukan pola pembelian pelanggan.

3.2. Pengumpulan Data

Pengumpulan data merupakan tahapan untuk mendapatkan data yang digunakan dalam melakukan penelitian setelah tahap perencanaan. Data yang digunakan yaitu data penjualan 212 Mart Jalan Rambutan Kota Pekanbaru sejak 21 Oktober 2018 sampai dengan 12 April 2019 sebanyak 13827 record. Adapun atribut yang digunakan adalah Jumlah Item Yang Dibeli dan Jumlah Item Yang Tidak Dibeli. Dua atribut ini merupakan acuan untuk pengelompokan produk, sedangkan atribut Tanggal Transaksi dan Item merupakan atribut nominal yang berfungsi pada saat analisa pola pengelompokan.

3.3. Pre-Processing

Tahap awal yang dilakukan ialah tahap pembersihan data, data yang akan dibersihkan adalah data transaksi pembelian produk selain produk/barang harian karena fokus dari penelitian ini yaitu produk/barang harian. Adapun atribut yang dilakukan proses transformasi yaitu atribut Item/Produk. Proses transformasi data dilakukan dengan cara membuat atribut baru yang mulanya hanya satu atribut umum yaitu "item/produk" menjadi atribut khusus nama-nama item/produk yang menjadi record data pada atribut sebelumnya. Langkah pertama yang dilakukan adalah memisahkan isi setiap data yang berisi lebih dari satu item, selanjutnya melakukan seleksi data yang dijadikan atribut-atribut baru. Langkah berikutnya melakukan pengisian data pada atribut-atribut baru dengan cara mengisi nilai 1 apabila terdapat item/produk sesuai atribut pada data dan mengisi nilai 0 apabila sebaliknya.

Atribut yang di normalisasi adalah atribut Jumlah Item Yang Dibeli dan Jumlah Item Yang Tidak Dibeli yang merupakan atribut numerik atau atribut kuantitatif yang bisa dihitung, adapun atribut Tanggal Transaksi dan Item merupakan atribut nominal dan tidak dinormalisasi. Berikut merupakan hasil dari tahap Pre-processing.

Tabel 1. Hasil Pre-Processing Data

No	Item Yang Dibeli	Item Yang Tidak Dibeli	41 JAHE MERAH 23gr	41 KOPI JAHE 23gr	...	ZWITSAL OIL NATURAL 100ml
1	4	3957	0	0	...	0
2	1	3960	0	0	...	0
3	2	3959	0	0	...	0
4	3	3958	0	0	...	0
...
13809	0,129	0,871	0	0	...	0

3.4. Pengelompokan Data Penjualan dengan Algoritma K-Medoids

Penerapan Metode K-Medoids dalam penelitian ini digunakan untuk pengelompokan data penjualan yang bertujuan agar dataset menjadi lebih kecil sehingga dapat membantu proses asosiasi menggunakan FP-

Growth nantinya. Pada penelitian ini untuk 13809 record data dilakukan percobaan jumlah cluster mulai dari 3 cluster hingga 9 cluster bertujuan untuk mendapatkan cluster terbaik. Data yang di inputkan pada setiap percobaan berjumlah 13809 record dengan jumlah maksimal iterasi 10. Berikut merupakan hasil pemrosesan data dengan RapidMiner pada setiap percobaan cluster:

1. Pada K=3, cluster 0 menghasilkan 11365 items, cluster 1 menghasilkan 1927 items dan cluster 2 menghasilkan 517 items.
2. Pada K=4, cluster 0 menghasilkan 11365 items, Cluster 1 menghasilkan 2342 items dan Cluster 2 menghasilkan 0 items dan Cluster 3 menghasilkan 102 items.
3. Pada K=5, Cluster 0 menghasilkan 8953 items, Cluster 1 menghasilkan 707 items dan Cluster 2 menghasilkan 3392 items dan Cluster 3 menghasilkan 757 items dan Cluster 4 menghasilkan 0 items.
4. Pada K=6, Cluster 0 menghasilkan 1116 items, Cluster 1 menghasilkan 6153 items dan Cluster 2 menghasilkan 757 items dan Cluster 3 menghasilkan 5783 items dan Cluster 4 menghasilkan 0 items dan Cluster 5 menghasilkan 0 items.
5. Pada K=7, Cluster 0 menghasilkan 1116 items, Cluster 1 menghasilkan 6153 items dan Cluster 2 menghasilkan 662 items dan Cluster 3 menghasilkan 5783 items dan Cluster 4 menghasilkan 0 items dan Cluster 5 menghasilkan 95 items dan Cluster 6 menghasilkan 0 items.
6. Pada K=8, Cluster 0 menghasilkan 1927 items, Cluster 1 menghasilkan 11365 items dan Cluster 2 menghasilkan 0 items dan Cluster 3 menghasilkan 431 items dan Cluster 4 menghasilkan 0 items dan Cluster 5 menghasilkan 0 items dan Cluster 6 menghasilkan 0 items dan Cluster 7 menghasilkan 86 items.
7. Pada K=9, Cluster 0 menghasilkan 3392 items, Cluster 1 menghasilkan 6153 items dan Cluster 2 menghasilkan 947 items dan Cluster 3 menghasilkan 431 items dan Cluster 4 menghasilkan 0 items dan Cluster 5 menghasilkan 2800 items dan Cluster 6 menghasilkan 0 items dan Cluster 7 menghasilkan 86 items dan Cluster 8 menghasilkan 0 items.

3.5. Uji Validasi Cluster

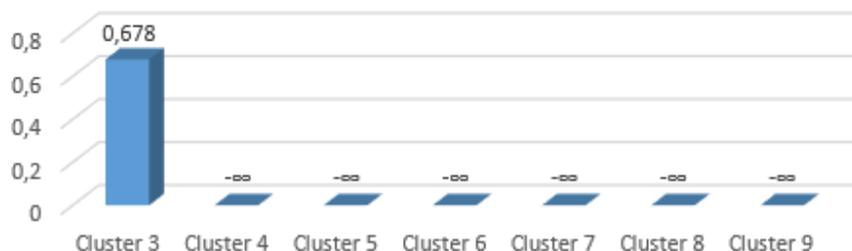
Setelah proses clustering selesai maka dilakukan validasi cluster menggunakan Davies Bouldien Index (DBI). Nilai DBI yang paling homogen adalah DBI dengan nilai yang paling kecil. Dari percobaan hasil 3 Cluster hingga 9 Cluster tersebut didapatkan nilai DBI pada setiap percobaan seperti yang terdapat pada tabel 2.

Tabel 2. Perbandingan Hasil Nilai DBI

Jumlah Cluster	Nilai DBI
cluster 3	0.678
cluster 4	-∞
cluster 5	-∞
cluster 6	-∞
cluster 7	-∞
cluster 8	-∞
cluster 9	-∞

Untuk lebih jelasnya, dapat dilihat pada gambar grafik 2 berikut.

Perbandingan Hasil Nilai DBI



Gambar 2. Grafik perbandingan hasil nilai DBI antar cluster

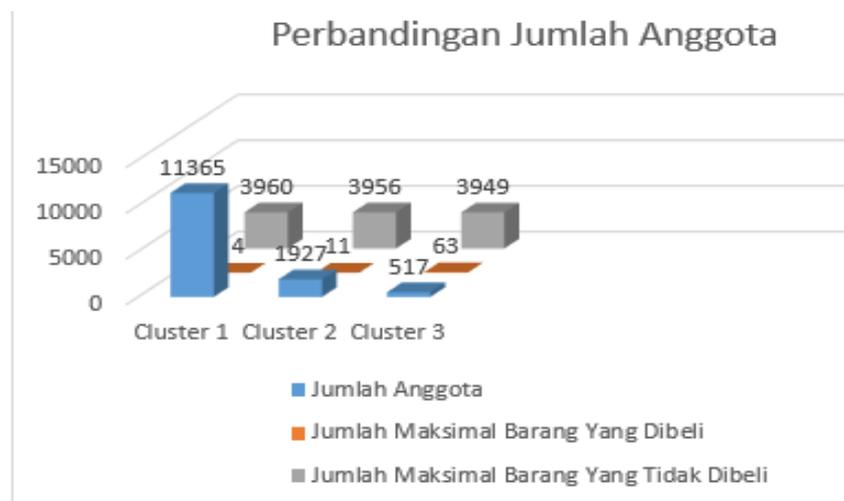
Berdasarkan gambar grafik 3 nilai DBI yang paling kecil adalah 3 Cluster dengan nilai DBI sebesar 0,678, maka data homogen jika dikelompokkan menjadi 3 kelompok. Informasi yang dapat diperoleh dari hasil clustering tersebut adalah pada transaksi penjualan yang semula terdiri dari 13809 record transaksi, dapat dikelompokkan menjadi 3 cluster:

1. Cluster pertama terdiri dari 11365 record transaksi, berisi transaksi yang memiliki jumlah barang yang dibeli sekitar 1-4, dan jumlah barang yang tidak dibeli sekitar 3957-3960.
2. Cluster kedua terdiri dari 1927 record transaksi, berisi transaksi yang memiliki jumlah barang yang dibeli sekitar 5-11, dan jumlah barang yang tidak dibeli sekitar 3950-3956.
3. Cluster ketiga terdiri dari 517 record transaksi, berisi transaksi yang memiliki jumlah barang yang dibeli sekitar 12-63, dan jumlah barang yang tidak dibeli sekitar 3898-3949.

Tabel 3. Hasil Cluster

Cluster	Jumlah Anggota	Jumlah Barang Yang Dibeli	Jumlah Barang Yang Tidak Dibeli
Cluster 1	11365	1 – 4	3957 – 3960
Cluster 2	1927	5 – 11	3950 – 3956
Cluster 3	517	12 – 63	3898 – 3949

Dan dapat dilihat pada gambar grafik 3 berikut jumlah anggota dari masing-masing cluster yang telah ditemukan.



Gambar 3. Grafik perbandingan pada cluster 3

3.6. Pencarian Association Rules Data Penjualan yang Telah di Cluster dengan Algoritma FP-Growth Menggunakan Tools RapidMiner

Setelah masing-masing dari cluster telah terbentuk, langkah selanjutnya adalah melakukan asosiasi menggunakan algoritma FP-Growth untuk menentukan rekomendasi produk kepada pelanggan. Hasil dari proses asosiasi ini diukur menggunakan nilai Support dan Confidence. Pada proses perhitungan menggunakan tools RapidMiner dilakukan percobaan dengan nilai minimal support 5% - 9% dan nilai Minimum Confidence 50%.

1. Asosiasi Cluster 1

Dari hasil pencarian rules/pola dengan menggunakan tools RapidMiner pada data cluster 1, dapat disimpulkan bahwa tidak ditemukan rules/pola dari nilai minimum support yang telah ditetapkan sebelumnya yaitu 5% sampai dengan 9%.

2. Asosiasi Cluster 2

Dari hasil pencarian rules/pola dengan menggunakan tools RapidMiner pada data cluster 2, dapat disimpulkan bahwa tidak ditemukan rules/pola dari nilai minimum support yang telah ditetapkan sebelumnya yaitu 5% sampai dengan 9%.

3. Asosiasi Cluster 3

Dari hasil pencarian rules/pola dengan menggunakan tools RapidMiner pada data cluster 3, dapat disimpulkan bahwa pada nilai minimum support sebesar 5% ditemukan rules/pola sebanyak 2 rules, pada minimum support sebesar 6% ditemukan rules/pola sebanyak 1 rules, pada minimum support sebesar 7% ditemukan rules/pola sebanyak 1 rules, pada nilai minimum support sebesar 8% ditemukan rules/pola sebanyak 1 rules dan pada nilai minimum support sebesar 9% tidak ditemukan rules/pola. Rules/pola dapat dilihat pada tabel 4.

Tabel 4. Hasil Association Rules Cluster 3

<i>Nilai Minimum Support</i>	<i>Premises</i>	<i>Conclusion</i>	<i>Support</i>	<i>Confidence</i>
5%	MINYAK GORENG PERMATA 2lt	GULA PASIR	0.052	0.563
6%	INDOMIE GORENG INDOMIE GORENG	INDOMIE KARI AYAM INDOMIE KARI AYAM	0.081 0.081	0.689 0.689
7%	INDOMIE GORENG	INDOMIE KARI AYAM	0.081	0.689
8%	INDOMIE GORENG	INDOMIE KARI AYAM	0.081	0.689
9%	-	-	-	-

3.7. Pencarian Association Rules Data Penjualan tanpa cluster dengan Algoritma FP-Growth Menggunakan Tools RapidMiner

Setelah pola asosiasi dari masing-masing cluster telah ditemukan, langkah selanjutnya adalah menemukan pola asosiasi data penjualan tanpa cluster menggunakan algoritma FP-Growth sehingga nantinya dapat dibandingkan rules mana yang lebih baik antara pola asosiasi data penjualan yang telah di cluster dengan pola asosiasi data penjualan tanpa cluster guna menentukan rekomendasi produk kepada pelanggan. Hasil dari proses asosiasi ini diukur menggunakan nilai Support dan Confidence. Pada proses perhitungan menggunakan tools RapidMiner dilakukan percobaan dengan nilai Minimal Support 5% - 9% dan nilai Minimum Confidence 50%. Dari hasil pencarian rules/pola dengan menggunakan tools RapidMiner pada data penjualan tanpa cluster, dapat disimpulkan bahwa tidak ditemukan rules/pola dari nilai minimum support yang telah ditetapkan sebelumnya yaitu 5% sampai dengan 9%.

3.8. Perbandingan Hasil Association Rules Setelah di Cluster Dengan Sebelum di Cluster

Dari hasil pencarian rules asosiasi dari data penjualan 212 Mart Jalan Rambutan Kota Pekanbaru yang telah dilakukan clustering terlebih dahulu didapatkan aturan asosiasi atau association rules pada cluster 3, yaitu sebanyak 5 rules, pada minimum support 5% ditemukan rules sebanyak 2 buah, pada minimum support 6% ditemukan rules sebanyak 1 buah, pada minimum support 7% ditemukan rules sebanyak 1 buah, dan pada minimum support 8% ditemukan rules sebanyak 1 buah.

Jika dibandingkan dengan data penjualan yang tanpa adanya proses cluster terlebih dahulu, rules sama sekali tidak dapat ditemukan. Hal ini sesuai dengan pendapat yang dikemukakan oleh Christidis dan kawan-kawan pada tahun 2010, bahwa masalah umum yang berhubungan dengan teknik rekomendasi, Dataset dengan jumlah yang besar cenderung diabaikan oleh aturan asosiasi. Sehingga menunjukkan bahwa dengan diimplementasikannya algoritma K-Medoids pada dataset dengan jumlah yang besar pada studi kasus ini dapat membantu dalam proses pencarian aturan asosiasi atau association rules guna mencari rekomendasi produk yang lebih akurat kepada pelanggan dikarenakan dataset yang akan diasosiasikan menjadi lebih kecil. Berdasarkan Association Rule yang telah terbentuk pada cluster 3 menyatakan bahwa ada 4 item/produk yang saling berhubungan yang paling sering dibeli oleh pelanggan 212 Mart Jalan Rambutan Kota Pekanbaru yaitu MINYAK GORENG PERMATA 2lt, GULA PASIR, INDOMIE GORENG dan INDOMIE KARI AYAM. Berdasarkan ketetapan nilai Support 5% - 9% dan Confidence 50%, berikut adalah rule-rule yang terbentuk.

1. Dengan nilai minimum support 5%, yaitu:
 - a. Jika pelanggan membeli produk MINYAK GORENG PERMATA 2lt maka dia juga akan membeli produk GULA PASIR dengan nilai keyakinan (Confidence) 56,3%.
 - b. Jika pelanggan membeli produk INDOMIE GORENG maka dia juga akan membeli produk INDOMIE KARI AYAM dengan nilai keyakinan (Confidence) 68,9%.
2. Dengan nilai minimum support 6%, yaitu:

Jika pelanggan membeli produk INDOMIE GORENG maka dia juga akan membeli produk INDOMIE KARI AYAM dengan nilai keyakinan (Confidence) 68,9%.
3. Dengan nilai minimum support 7%, yaitu:

Jika pelanggan membeli produk INDOMIE GORENG maka dia juga akan membeli produk INDOMIE KARI AYAM dengan nilai keyakinan (Confidence) 68,9%.
4. Dengan nilai minimum support 8%, yaitu:

Jika pelanggan membeli produk INDOMIE GORENG maka dia juga akan membeli produk INDOMIE KARI AYAM dengan nilai keyakinan (Confidence) 68,9%.

4. KESIMPULAN

Adapun kesimpulan dari penelitian ini adalah sebagai berikut:

1. Setelah melakukan beberapa percobaan dengan jumlah cluster 3, 4, 5, 6, 7, 8 dan 9 sebagai permodelan percobaan dengan menggunakan algoritma K-Medoids clustering, didapatkanlah cluster optimal sebanyak 3 cluster berdasarkan uji validitas Davies Bouldien Indeks (DBI) dengan nilai sebesar 0,678.
2. Berdasarkan percobaan pencarian Association Rule dari 3 cluster dengan nilai minimum support sebesar 5% - 9% dan nilai minimum confidence 50% didapatkanlah hasil bahwa Association Rule ditemukan hanya pada cluster 3.
3. Association Rule yang terbentuk pada cluster 3 dengan nilai minimum support sebesar 5% ditemukan rules/pola sebanyak 2 rules, pada minimum support sebesar 6% ditemukan rules/pola sebanyak 1 rules, pada minimum support sebesar 7% ditemukan rules/pola sebanyak 1 rules, pada minimum support sebesar 8% ditemukan rules/pola sebanyak 1 rules dan pada minimum support sebesar 9% tidak ditemukan rules/pola.
4. Berdasarkan Association Rule yang telah terbentuk pada cluster 3 menyatakan bahwa ada 4 item/produk yang saling berhubungan yang paling sering dibeli oleh pelanggan 212 Mart Jalan Rambutan Kota Pekanbaru yaitu MINYAK GORENG PERMATA 2lt, GULA PASIR, INDOMIE GORENG dan INDOMIE KARI AYAM yang dapat direkomendasikan kepada pihak 212 Mart.

REFERENSI

- [1] Aguinis, Herman, Lura E. Forcum, dan Harry Joo. Using Market Basket Analysis in Management Research. *Journal of Management*. 2012; 39(7): 1799-1824.
- [2] Annie, Loraine Charlet, dan Ashok Kumar. Market Basket Analysis for a Supermarket based on Frequent Itemset Mining. *International Journal of Computer Science*. 2012; 5(9): 257-264.
- [3] Ansari, Zahid., dkk. Quantitative Evaluation of Performance and Validity Indices for Clustering the Web Navigational Sessions. *World of Computer Science and Information Technology Journal (WCSIT)*. 2011; 1(5): 217-226.
- [4] Azhari, Anshori. Pendekatan Aturan Asosiasi Untuk Analisis Pergerakan Saham. Seminar Nasional Informatika. Yogyakarta. 2009; 1, 183-189.
- [5] Bhatia, Shveta K, V.S Dixit. A Propound Method for the Improvement of Cluster Quality. *International Journal of Computer Science Issues*. 2012; 9(4): 216-222.
- [6] Christidis, Konstantinos, Dimitris Apostolou, dan Gregoris Mentzas. Exploring Customer Preferences with Probabilistic Topics Models, 1-13.
- [7] Diwate Rahul B, Sahu Amit. Data Mining Techniques in Association Rule : A Review. *International Journal of Computer Science and Information Technologies*. 2014; 5(1): 227-229.
- [8] Erwin. Analisis Market Basket Dengan Algoritma Apriori dan FP-Growth. *Jurnal Generic*. 2009; 2(4): 26-30.
- [9] Fitria, Rizky, Warnia Nengsih, Dini Hidayatul Qudsi. Implementasi Algoritma FP-Growth Dalam Penentuan Pola Hubungan Kecelakaan Lalu Lintas. *Jurnal Sistem Informasi*. 2017; 13 (2): 118-124.
- [10] Ghuman, Sukhdev Singh. Clustering Techniques- A Review. *International Journal of Computer Science and Mobile Computing*. 2016; 5(5): 524-530.
- [11] Gupta, Savi, dan Roopal Mamtora. A Survey on Association Rule Mining in Market Basket Analysis. *International Journal of Information and Computation Technology*. 2014; 4(4): 409-414.
- [12] Kaur, Noor Kamal, Usvir Kaur, dan Dr Dheerendra Singh. K-Medoid Clustering Algorithm- A Review. *International Journal of Computer Application and Technology*. 2014; 1 (1): 42-45.
- [13] Liu, Xi, Xiumei Zhang, Xiyong Li, dan Zhenguang Sun. Research on Data Mining Clustering Algorithm in Cloud Computing Environments. *BioTechnology An Indian Journal*. 2014; 10(17): 9563-9566.
- [14] Kamila I, Khairunnisa U, Mustakim M, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau", *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, Vol 5 No 1, pp: 119-125, 2019.
- [15] Mustakim M, Della Maulina Herianda, Ahmad Ilham, Achmad Daeng GS, Folkes E. Laumal, Nuning Kurniasih, Akbar Iskandar, Gloria Manulangga, Ida Bagus Ary Indra Iswara, dan Robbi Rahim, "Market Basket Analysis Using Apriori and FP-Growth for Analysis Consumer Expenditure Patterns at Berkah Mart in Pekanbaru Riau", *Journal of Physics*, Vol. 1114 No. 1, pp: 1-9, 2018.
- [16] Park, Hae-Sang, Jong-Seok Lee, dan Chi-Hyuck Jun. A K-means-like Algorithm for K-medoids Clustering and Its Performance. 2014.
- [17] Patel, Dhara, Ruchi Modi, dan Ketan Sarvakar. A Comparative Study of Clustering Data Mining: Techniques and Research Challenges. *International Journal of Latest Technology in Engineering, Management & Applied Science*. 2014; 3(9): 67-70.

- [18] Plasse, Marie, Ndeye Niang, Gilbert Saporta, Alexandre Villeminot, dan Laurent Leblond. Combined Use of Association Rules Mining and Clustering Methods to Find Relevant Links between Binary Rare Attributes in a Large Data Set. *Computational Statistics & Data Analysis*. 2007; 52(1): 596-613.
- [19] Rahayu G, dan Mustakim M, "Principal Component Analysis untuk Dimensi Reduksi Data Clustering Sebagai Pemetaan Persentase Sertifikasi Guru di Indonesia", Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI), Vol. 9, pp: 201-208, 2017.
- [20] Rani, Yogita, dan Dr Harish Rohil. A Study of Hierarchical Clustering Algorithm. *International Journal of Information and Computation Technology*. 2013; 3(11): 1225-1232.
- [21] Raval, Unnati R. Implementing & Improvisation of K-Means Clustering Algorithm. *International Journal of Computer Science and Mobile Computing*. 2016; 5(5): 191-203.
- [22] Sinthuja, M, Dr N Puviarasan, Dr P Aruna. Research of Improved FP-Growth (IFP) Algorithm in Association Rules Mining. *National Conference on Internet of Things IOT*. 2018; 4: 24-31.
- [23] Triyanto, Wiwit Agus, Vincent Suhartono, dan Himawan. Analisis Keranjang Belanja Menggunakan K-Medoids dan FP-Growth. *Jurnal Pseudocode*. 2014; 2(1): 129-142.
- [24] Velmurugan, Dr T. 2012. "Efficiency of K-Means and K-Medoids Algorithms for Clustering Arbitrary Data Points" 3: 7.
- [25] Vijayarani, Dr S, dan R Prasannalakshmi. Association Rule Generation in Data Streams Using FP-Growth and APRIORI MR Algorithms. *International Journal of Innovative Research in Computer and Communication Engineering*. 2015; 3(9): 8949-8956.
- [26] Wen-xiu, Xie, Qi Heng-nian, dan Huang Mei-li. Market Basket Analysis Based on Text Segmentation and Association Rule Mining. *First International Conference on Networking and Distributed Computing*. 2010. 309-313.

BIBLIOGRAFI PENULIS



Imaduddin Syukra, merupakan salah satu alumni dari Universitas Islam Negeri Sultan Syarif Kasim Riau, program studi Sistem Informasi. Dia bekerja sebagai Back-End Developer di MENTE I.C.E ROBOTIC yang memiliki ketertarikan di bidang Big Data khususnya Asosiasi. Semasa kuliah bergabung di salah satu organisasi di Fakultas Sains dan Teknologi yakni di Puzzle Research Data Technology (PREDATECH).



Assad Hidayat, merupakan salah satu alumni dari Universitas Islam Negeri Sultan Syarif Kasim Riau, program studi Sistem Informasi. Dia bekerja sebagai IT System Trainer dan memiliki ketertarikan di bidang Machine Learning dan Big Data. Semasa kuliah pernah bergabung di salah satu organisasi di Fakultas Sains dan Teknologi yaitu Puzzle Research Data Technology (PREDATECH).



Muhammad Zakiy Fauzi, merupakan salah satu mahasiswa Universitas Islam Negeri Sultan Syarif Kasim Riau, program studi Sistem Informasi. Memiliki ketertarikan di bidang Big Data khususnya Data Mining dan Text Mining. Semasa kuliah bergabung di salah satu organisasi di Fakultas Sains dan Teknologi yakni di Puzzle Research Data Technology (PREDATECH).