

Paper

Data Mining Market Basket Analysis Menggunakan Algoritma Hash Based Pada Sistem Penjualan Produk

Author : M Fadli Hamdi Lubis



Program Studi Teknik Informatika

TEMA : DUNIA AKADEMIK DI ERA SOCIETY 5.0 : TANTANGAN DAN PELUANG



ISSN : 2964-1950

Data Mining Market Basket Analysis Menggunakan Algoritma Hash Based Pada Sistem Penjualan Produk

M Fadli Hamdi Lubis

Universitas Harapan Medan, Medan, Indonesia

¹fadlihamdi63@gmail.com,

Abstrak

Penelitian ini bertujuan untuk menentukan pola belanja konsumen dengan Data Mining *Market Basket Analysis* menggunakan Algoritma *Hash Based* pada sistem penjualan produk di dapat hasil *item* yang paling banyak dibeli berdasarkan periode Agustus 2022 – Juli 2023 untuk 1-*itemset* adalah Sanger Ice (A) Sebesar 42%. Untuk 2-*itemset* yaitu Sanger Ice (A) dan Sanger Hot (B) Sebesar 25%. Setelah dilakukan tiga kali *iterasi* tersisa satu alamat pada tabel *hash* dengan 3-*itemset*. *Frequent itemset* yang di dapat yaitu Sanger Ice (A), Sanger Hot (B), Sarang Burung (AO) dengan jumlah yang terjual 42 (*Support Count*) atau 12% Dan *Confidence* 49%.

Kata Kunci: Data Mining, Market Basket Analysis, Algoritma Hash Based, Itemset

Abstract

This research aims to determine consumer shopping patterns with Data Mining Market Basket Analysis using the Hash Based Algorithm in the product sales system. The results show that the most purchased item based on the period August 2022 - July 2023 for 1-itemset is Sanger Ice (A) at 42%. For 2-itemset, namely Sanger Ice (A) and Sanger Hot (B), it is 25%. After three iterations, one address remains in the 3-itemset hash table. Frequent itemsets obtained are Sanger Ice (A), Sanger Hot (B), Bird's Nest (AO) with a quantity sold of 42 (Support Count) or 12% and Confidence 49%

Keywords: Data Mining, Market Basket Analysis, Algorithm Hash Based, Itemset

1. PENDAHULUAN

Kopi Masjid Abidin merupakan sebuah usaha yang berdiri sejak tahun 2020. Kopi ini berlokasi di Jl Brigjend Katamsa Medan. Kopi Masjid ini menjual minuman kopi dan makanan ringan pada umumnya. Kopi Masjid masih menggunakan sistem manual pada setiap data transaksi yaitu dicatat dan hanya diarsipkan begitu saja. Kopi Masjid tidak memanfaatkan data yang sudah ada untuk menyusun atau menentukan strategi pemasaran. Berdasarkan temuan pasca Kopi Masjid tersebut membutuhkan suatu sistem agar data yang diperoleh dapat dijadikan untuk informasi yang berguna dalam mengembangkan usaha tersebut Kopi Masjid dapat menentukan salah satu strategi pemasaran untuk meningkatkan penjualan yaitu dengan menggunakan data-data transaksi yang sudah ada sebelumnya kemudian dijadikan pertimbangan untuk mendapatkan produk rekomendasi berdasarkan *item* pemesanan yang dipesan konsumen. Adapun untuk mengolah data-data transaksi tersebut menjadi sebuah informasi atau pengetahuan diperlukan suatu teknik atau metode yaitu dengan *data mining*. *Data Mining* adalah bidang ilmu komputer yang dapat memberikan kemampuan mengekstrak berbagai data yang banyak dan kompleks guna memperoleh wawasan yang berharga dari berbagai macam data. pada konteks penjualan produk pada Masjid Kopi, *Data Mining* mampu menjadi solusi yang bagus untuk melakukan analisis pasar berbasis data[1]. Dengan menganalisis data penjualan dan perilaku konsumen, perusahaan dapat melakukan *mengidentifikasi* tren pasar, memahami *preferensi* konsumen, dan *mengoptimalkan* strategi penjualan produk. Dalam menemukan pola dari data-data transaksi tersebut maka perlu dianalisis dengan menggunakan *Market Basket Analysis*. *Market Basket Analysis* (analisis keranjang belanja) atau yang biasa dikenal *Association Rules Mining* adalah teknik *data mining* yang berasal dari bidang pemasaran. Teknik ini digunakan untuk menentukan produk-produk manakah yang akan dibeli oleh konsumen secara bersamaan dengan melakukan analisis terhadap daftar transaksi konsumen[2]. Algoritma *Hash Based* merupakan algoritma yang bertujuan untuk menemukan pola kombinasi dari beberapa alternatif dengan kriteria atau atribut yang dijadikan sebagai pembanding dengan perhitungan menggunakan teknik penyaringan sehingga dapat menghasilkan suatu pola kombinasi *itemset*. Algoritma ini merupakan algoritma yang dikembangkan dari algoritma apriori, sehingga fungsi algoritma ini sama dengan fungsi algoritma apriori[3]. Pola kombinasi yang diperoleh dari pengelolaan data-data hasil transaksi yang telah dilakukan pada penjualan berbagai *itemset* sebelumnya. Sehingga dalam pengelolaan data menggunakan algoritma *Hash Based* ini harus dilakukan dengan menggunakan data dari hasil transaksi hasil penjualan sebelumnya. Melihat pentingnya analisis pasar berbasis data dan keunggulan algoritma *hash based* dalam pemrosesan data, penelitian ini bertujuan untuk mengaplikasikan *Data Mining Market Basket Analysis* menggunakan algoritma *Hash Based* pada sistem penjualan produk di kopi masjid. Penelitian ini diharapkan dapat memberikan pemahaman yang lebih baik tentang perilaku pasar dan *preferensi* konsumen, serta memberikan

wawasan yang berguna bagi perusahaan dalam mengambil keputusan strategis dalam penjualan dan pemasaran produk. Dengan memahami tren pasar, *preferensi* konsumen, dan pola penjualan produk, perusahaan dapat *mengoptimalkan* strategi penjualan, *mengidentifikasi* peluang baru, dan meningkatkan kepuasan konsumen [4]. Selain itu, hasil penelitian ini dapat menjadi dasar bagi pengembangan sistem informasi penjualan yang lebih efisien dan efektif di Kopi Masjid.

Dalam penelitian dengan judul Penerapan *Data Mining* Algoritma *Hash Based* Pada Data Pemesanan Buah Impor Cv Green Uni Fruit mendapatkan hasil berupa sebuah pola asosiasi buah jeruk sunkkist, anggur sunrale dan leci merupakan kombinasi buah yang diprioritaskan untuk pemesanan dengan nilai *support* sebesar 25% dan *confidence* 60% [5]. Penelitian lainnya yang berjudul Analisis *Market Basket* Dengan Algoritma *Hash-Based* Pada Transaksi Penjualan (Studi Kasus: Tb. Menara) didapatkan hasil penelitian Metode *Market Basket Analysis* dengan algoritma *hash-based* dapat digunakan untuk membantu toko mengetahui pola belanja konsumen dan *item* yang sering dibeli oleh konsumen [6]. Penelitian yang dilakukan oleh dengan judul Penemuan Pola Asosiasi Pada Data Restoran Menggunakan Algoritma *Hash Based* mendapatkan hasil penelitian dengan menggunakan algoritma *hash based* dapat digunakan untuk menentukan kombinasi pola *itemset* [7].

2. METODE PENELITIAN

2.1 Metodologi Penelitian

Dalam metodologi penelitian ini terdapat metode pengumpulan data dan metode analisis data sebagai berikut:

2.1.1 Metode Pengumpulan Data

Dalam penelitian ini untuk mendapatkan suatu data dan informasi penulis memerlukan suatu metode yang akan digunakan untuk pengumpulan data. metode yang dilakukan penulis menggunakan metode pengumpulan data yang terdiri dari tiga macam teknik pengumpulan data:

1. Interview (Wawancara)

Teknik wawancara atau *interview* adalah cara untuk memperoleh informasi secara langsung secara tatap muka melalui tanya jawab antara pengumpul data dengan *responden* atau *narasumber*. Teknik pengumpulan data dengan wawancara biasanya dilakukan sebagai studi pendahuluan, karena teknik ini tidak mungkin dilakukan jika *respondennya* dalam jumlah besar.

2. Observasi (Pengamatan)

Teknik *Observasi* atau Pengamatan merupakan cara yang dilakukan *responden* terjun langsung ke tempat untuk mendapatkan informasi-informasi yang penting. Dalam penelitian ini penulis melakukan pengamatan terhadap objek data dengan mendapatkan *korelasi* antara data transaksi penjualan dengan kebutuhan data yang akan digunakan.

3. Studi Pustaka

Teknik studi pustaka yaitu dengan cara yang dilakukan dengan pengumpulan data yang memanfaatkan berbagai hal atau variabel. Teknik pengumpulan data dengan dokumentasi biasanya dilakukan dengan memperoleh data dan informasi dalam bentuk buku, jurnal, dokumen, artikel dan arsip yang berkaitan dengan penelitian.

2.1.2 Metode Analisis Data

Adapun metode yang digunakan didalam penerapan *data mining* adalah menggunakan proses tahapan KDD (*Knowledge Discovery in Database*) yang terdiri dari beberapa tahapan seperti berikut ini:

1. Pemilihan Data (*Data Selection*), pemilihan data relevan yang di dapat dari buku.

2. Pembersihan Data (*Data Cleaning*), proses penghapusan data yang tidak penting atau tidak relevan.

3. Transformasi Data (*Data Transformation*), data diubah ke dalam format yang sesuai untuk diproses dalam *data mining*.

4. *Data Mining*, suatu proses dimana metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

5. *Interpretation / Evaluation*, *visualisasi* dan penyajian pengetahuan teknik yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

2.2 Market Basket Analysis

Market Basket Analysis adalah suatu metodologi untuk melakukan analisis *buying habit* konsumen dengan menemukan asosiasi antar beberapa *item* yang berbeda, yang diletakkan konsumen dalam *shopping basket* (keranjang belanja) yang dibeli pada suatu transaksi tertentu. Tujuan dari *market basket analysis* adalah untuk mengetahui produk-produk mana yang mungkin akan dibeli secara bersamaan [8]. Pihak *retailer* dapat

menggunakan informasi ini untuk menempatkan barang-barang yang sering terjual bersama di area yang sama[9].

2.3 Association Rules Mining

Association Rules Mining adalah suatu prosedur untuk mencari hubungan antar *item* dalam suatu *dataset*. *Association rules* meliputi dua tahap[10]. antara lain sebagai berikut:

- 1) Mencari kombinasi yang paling sering terjadi dari suatu *itemset*.
- 2) Mendefinisikan *Condition* dan *Result* (untuk *Conditional Association Rules*). Dalam menentukan suatu *Association Rules*, terdapat suatu *Interestingness Measure* (ukuran kepercayaan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu.

Rumus dalam mencari nilai *minimum support* sebuah *item*:

$$\text{Support}(X) = \frac{\text{Jumlah transaksi mengandung } (X)}{\text{Total transaksi}} \quad (2.1)$$

Sedangkan untuk mencari nilai *support* untuk *itemset* (X,Y):

$$\text{Support}(X,Y) = \frac{\text{Jumlah transaksi mengandung } X \text{ dan } Y}{\text{Total transaksi}} \quad (2.2)$$

Rumus dalam mencari nilai *minimum confidence*:

$$\text{Confidence}(Y|X) = \frac{\text{Jumlah transaksi mengandung } X \text{ dan } Y}{\text{Jumlah transaksi } X} \quad (2.3)$$

2.4 Algoritma Hash Based

Algoritma *Hash Based* menggunakan teknik *hashing* untuk menyaring keluar *itemset* yang tidak penting untuk pembangkitan *itemset* selanjutnya. Ketika *support count* untuk kandidat *k-itemset* dihitung dengan menelusuri *database*, algoritma *hash-based* mengumpulkan informasi mengenai *(k+1)-itemset* dengan cara seluruh kemungkinan *(k+1)-itemset* di *hash* kedalam *hash table* dengan menggunakan fungsi *hash* (yang menggunakan sebuah bilangan prima untuk operasi modulo)[5].

Pada jurnal “*A Hash Based Frequent Itemset Mining using Rehashing*”, Aguru,dkk melakukan pengurangan ukuran dengan menggunakan *hash function* pada level dua dan tiga.

Berikut adalah rumus *hash* pada level kedua dan ketiga:

Rumus *Hash* untuk pemrosesan *hash table* pada level kedua

$$H\{XY\} = (\text{order of } X) * 10 + (\text{order of } Y) \text{ mod bil prima} \quad (2.4)$$

Rumus *Hash* untuk pemrosesan *hash table* pada level ketiga

$$H\{XYZ\} = (\text{order of } X * 100) + (\text{order of } Y * 10) + \text{order of } Z \text{ mod bil prima} \quad (2.5)$$

Keterangan:

Order of X = perwakilan nilai X Order of Y = perwakilan nilai Y

Order of Z = perwakilan nilai Z H = alamat pada tabel *hash*

n = banyak alamat (n = 2m+1, m = jumlah keseluruhan *item*)

Mod bil prima = bilangan prima yang terdekat dan yang lebih besar dari jumlah kombinasi (kombinasi *k-itemset*).

3. HASIL DAN PEMBAHASAN

3.2 Perancangan Sistem

Dalam perancangan sistem ini dibuat dengan algoritma *hash based* untuk analisis data penjualan data transaksi sebanyak 340 transaksi dalam waktu satu tahun.

Tabel 1. Data Transaksi Penjualan

No Transaksi	Item Pemesanan
1	Sanger Ice,Sanger Hot,Espresso Muslimin,Teh Kotak,Americano,Black Ice Drip,Teh Manis Ice
2	Sanger Ice,Sanger Al Ikhlas,Black Ice Drip,Sarang Burung,Golda Coffee,BearBrands,Mizone,Teh Manis Ice
3	Sanger Ice,Sanger Al Ikhlas,Jangek,Fresstea,Sanger Ice Milk
...
340	Sarang Burung,Golda Coffee,Pulpy Orange,Sanger Ice,Fruit Tea,Teh Pucuk

3.1 Analisis Algoritma Hash Based

Dalam proses algoritma *hash based* langkah yang dilakukan adalah menentukan *minimum support (minsup)* sebagai batas pembentukan *frequent itemset* dan *minimum confidence (mincof)* setelah dilakukan *preprocessing* data. *Preprocessing* data dilakukan untuk memproses kembali data tersebut seperti ada nilai yang hilang, data yang berlebih atau *format* yang tidak sesuai. penentuan *minimum support* dan *minimum confidence* disesuaikan dengan kebutuhan karena tidak ada ketentuannya. Dalam penelitian ini peneliti menggunakan *minimum support* sebesar 63 atau 19% dan *minimum confidence* 40%.

3.2 Proses Algoritma Hash Based

Untuk mempermudah dalam proses algoritma *hash based* perhitungan tabel *hash* dilakukan urutan data setiap *item* yang digunakan menggunakan inisial abjad

Tabel 2. Inisial Item

No	Inisial Item	Nama Item	Support
1	A	Sanger Ice	42%
2	B	Sanger Hot	41%
3	C	Sanger Ice Milk	17%
...
48	AV	Floridina	1%

Untuk perhitungan support setiap *item* dilakukan dengan rumus (2.1):

$$\text{Support (X)} = \frac{\text{Jumlah Transaksi Mengandung X}}{\text{Total Transaksi}}$$

$$\text{Sanger Ice} = \frac{144}{340} = 42\%$$

$$\text{Sanger Ice Milk} = \frac{58}{340} = 17\%$$

$$\text{Sanger Hot} = \frac{139}{340} = 41\%$$

$$\text{Floridina} = \frac{3}{340} = 1\%$$

Hasil perhitungan C1 tersebut kemudian disaring berdasarkan nilai *support count* yang lebih besar atau sama dengan dari nilai *minimum support* yang telah ditentukan sebelumnya yaitu sebesar 19%. Hasil dari penyaringan tersebut menghasilkan 1-*Frequent itemset* (L1). Hasil penyaringan L1 dapat dilihat pada tabel di bawah ini

Tabel 3. Frequent Itemset (L1)

No	Item	Support
1	Sanger Ice (A)	42%
2	Sanger Hot (B)	41%
3	Sanger Al Ikhlas (D)	30%
4	Espresso Muslimin (E)	20%
5	Americano (AJ)	23%
6	Golda Coffee (AK)	26%

7	Sarang Burung (AO)	34%
---	--------------------	-----

Hasil L1 dikombinasikan dan dimasukkan ke dalam tabel *hash* dengan menggunakan persamaan 2.4.
 $H\{XY\} = (\text{order of } X) * 10 + (\text{order of } Y) \text{ mod bil prima}$

Tabel 4. Perhitungan Tabel Hash Alamat 2-Itemset

No	Kombinasi 2-Itemset	Perhitungan	Hasil
1	A,B	$[(1) * 10 + (2)] \text{ mod } 23$	12*
2	A,D	$[(1) * 10 + (3)] \text{ mod } 23$	13*
3	A,E	$[(1) * 10 + (4)] \text{ mod } 23$	14*
...
21	AK,AO	$[(6) * 10 + (7)] \text{ mod } 23$	21

Pada perhitungan *hashing* tersebut ditemukan adanya *collision* (Adanya lebih dari 1 *itemset* yang memiliki alamat *hash* sama). Pada kasus ini *collision* terjadi pada alamat ke-0 untuk (B,D) dengan (E,AK), alamat ke-1 untuk (B,E) dengan (E,AO), alamat ke-11 untuk (D,E) dengan (AJ,AO), alamat ke-12 untuk (A,B) dengan (D,AJ). alamat ke-13 untuk (A,D) dengan (D,AK), alamat ke-14 untuk (A,E) dengan (D,AO),
 Ketika sebuah *collision* terjadi maka harus segera dilakukan pengecekan untuk mencari alamat *bucket* yang masih kosong. Jika setelah dilakukan pengecekan ditemukan *indikasi* bahwa tabel *hash* sudah separuhnya terisi maka perlu dilakukan *rehashing* ulang dengan banyak alamat 2 kali dari banyak alamat sebelumnya. Untuk menyelesaikan masalah tersebut digunakan rumus yang berbeda yaitu:

$$H(K) = ((\text{Order of } X) * 10 + \text{Order of } Y) \text{ mod } n$$

$n =$ banyak alamat setelah dilakukan penambahan ($n = 2 * m + 1$, $m =$ jumlah alamat pada tabel *hash* sebelum dilakukan penambahan)

Tabel 5. Perhitungan Tabel Hash Setelah Penambahan Alamat

No	Kombinasi 2-Itemset	Perhitungan	Hasil
1	A,B	$[(1) * 10 + (2)] \text{ mod } 47$	12
2	A,D	$[(1) * 10 + (3)] \text{ mod } 47$	13
3	A,E	$[(1) * 10 + (4)] \text{ mod } 47$	14
...
21	AK,AO	$[(6) * 10 + (7)] \text{ mod } 47$	20

Setelah dilakukan penambahan alamat tabel *hash*, tidak ditemukan lagi adanya *collision* antar *itemset*. Setiap alamat terisi dengan 1 *itemset*.

Untuk pembangkitan L2 di dapat berdasarkan perhitungan *support count subset* C2 dengan menggunakan Rumus 2.2. C2 yang tidak memiliki *support count* lebih besar atau sama dengan dari *minimum support* tidak akan dibangkitkan menjadi L2.

$$\text{Support } (X, Y) = \frac{\text{Jumlah Transaksi Mengandung } X \text{ dan } Y}{\text{Total Transaksi}}$$

$$\text{Support } (A, B) = \frac{86}{340} = 25\%$$

$$\text{Support } (A, E) = \frac{28}{340} = 8\%$$

$$\text{Support } (A, D) = \frac{6}{340} = 2\%$$

$$\text{Support } (AK, AO) = \frac{39}{340} = 11\%$$

Dari perhitungan *support count* diatas C2 yang memiliki nilai *support count* lebih besar atau sama dengan 19% maka akan dibangkitkan menjadi L2. L2 yang terbentuk dapat dilihat pada tabel di bawah ini

Tabel 6. Frequent Itemset (L2)

No	Alamat	Item	Nama Item	Support	Confidence
1	12	A,B	Sanger Ice - Sanger Hot	25%	60%
2	17	A,AO	Sanger Ice - Sarang Burung	19%	45%
3	27	B,AO	Sanger Hot - Sarang Burung	19%	46%

Setelah ditemukan L2 maka L2 akan dikombinasikan dan di *hash* ke dalam *hash table* dengan persamaan 2.5. Hasil L2 yang telah dikombinasikan dimasukkan ke dalam *bucket address*.

$H\{XYZ\} = (\text{order of } X * 100) + (\text{order of } Y * 10) + \text{order of } Z \text{ mod bil prima}$

Tabel 7. Perhitungan Alamat Tabel *Hash 3-Itemset*

No	Kombinasi 3-Itemset	Perhitungan	Hasil
1	A,B,AO	$[(1) * 100 + (2) * 10 + 3] \text{ mod } 47$	29

Tabel 8. Frequent Itemset (L3)

No	Kombinasi 3-Itemset	Nama Item	Support Count	Support	Confidence
1	A,B \Leftrightarrow AO	Sanger Ice, Sanger Hot, Sarang Burung	42	12%	49%

Setelah dilakukan tiga kali *iterasi* tersisa satu alamat pada tabel *hash* dengan *3-itemset*. *Frequent itemset* yang didapatkan yaitu (Sanger Ice, Sanger Hot, Sarang Burung) dengan jumlah *item* yang laku terjual secara bersamaan sebesar 42.

Setelah tahap perhitungan algoritma *hash-based* selesai maka dilakukan perhitungan *confidence*. Penyaringan hasil perhitungan *confidence* dilakukan berdasarkan *minimum confidence* yang telah ditetapkan dengan menggunakan Rumus 2.3 untuk semua hasil perhitungan algoritma *hash-based*.

$$\text{Confidence (Y|X)} = \frac{\text{Jumlah Transaksi Mengandung X dan Y}}{\text{Jumlah Transaksi X}}$$

$$\text{Confidence (B|A)} = \frac{86}{144} = 60\%$$

$$\text{Confidence (AO|A)} = \frac{65}{144} = 45\%$$

$$\text{Confidence (AO|B)} = \frac{64}{139} = 46\%$$

$$\text{Confidence (AO|A, B)} = \frac{42}{84} = 49\%$$

Setelah di dapat nilai *support* dan *confidence* di atas maka akan dibentuk sebuah *rules* yang memenuhi *minimum support* 40% dengan *rules* yang terbentuk yaitu:

[Sanger Ice (A)] – [Sanger Hot (B)] = Confidence 60%

[Sanger Ice (A)] – [Sarang Burung (AO)] = Confidence 45%

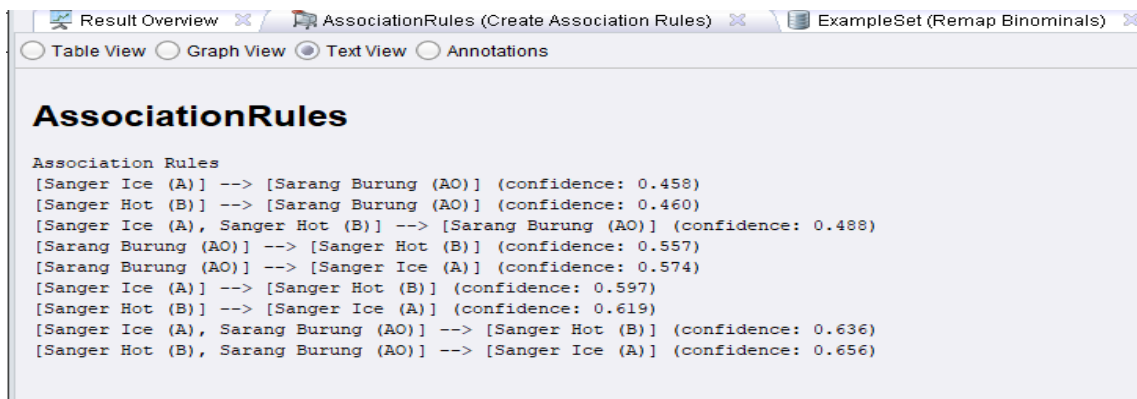
[Sanger Hot (B)] – [Sarang Burung (AO)] = Confidence 46%

[Sanger Ice (A) - Sanger Hot (B)] – [Sarang Burung (AO)] = Confidence 49%

Untuk mengimplementasikan pengujian sistem yang ingin dicari oleh penulis, penulis menggunakan *tools rapidminer 5.2* untuk sebagai alat untuk menerapkan algoritma *hash based* dengan *minimum support* 19% dan *minimum confidence* 40%.

3.3 Implementasi Menggunakan Rapidminer 5.2

Disini peneliti menggunakan file CSV dari data yang di *tabular* setiap transaksi selama satu tahun 340 Transaksi. Pembentukan dalam proses algoritma *hash based* pada *rapidminer 5.2*



Gambar 1. Hasil Association Rules Text View

Dari hasil *association rules* dalam bentuk *text view* dapat dilihat *confidence* dari perhitungan menggunakan aturan asosiasi pada *rapidminer* menghasilkan:

- jika membeli Sanger Ice (A), maka membeli Sanger Hot (B) dengan *confidence* 0.597 (60%).
- Jika membeli Sanger Ice (A) maka membeli Sarang Burung (AO) dengan *confidence* 0.458 (45%).
- jika membeli Sanger Hot (B) maka membeli Sarang Burung (AO) dengan *confidence* 0.460 (46%) dan
- jika membeli Sanger Ice (A), Sanger Hot (B) maka membeli Sarang Burung (AO) dengan *confidence* 0.488 (49%).

4. KESIMPULAN

Berdasarkan penelitian yang dilakukan penulis bahwa:

- Dengan menentukan pola belanja konsumen yang terjadi menggunakan algoritma *hash-based* dan penerapannya dalam *Data Mining Market Basket* menggunakan algoritma *Hash Based* pada sistem penjualan produk menggunakan *tools rapidminer 5.2* maupun Microsoft Excel dalam perhitungan manual menghasilkan aturan yang sama.
- Dengan aturan yang dihasilkan dari penerapan algoritma *hash based* dengan nilai *minimum support* 19% dan *minimum confidence* 40% sebagai berikut:
 - Jika membeli Sanger Ice (A) maka akan terbeli bersama dengan Sanger Hot (B) dengan tingkat kepercayaan 60% dan didukung oleh 25% data keseluruhan.
 - Jika membeli Sanger Ice (A) maka akan terbeli bersama dengan Sarang Burung (AO) dengan tingkat kepercayaan 45% dan didukung oleh 19% data keseluruhan.
 - Jika membeli Sanger Hot (B) maka akan terbeli bersama dengan Sarang Burung (AO) dengan tingkat kepercayaan 60% dan didukung oleh 19% data keseluruhan.
 - Jika membeli Sanger Ice (A) dan Sanger Hot (B) maka akan terbeli bersama dengan Sarang Burung (AO) dengan tingkat kepercayaan 49% dan didukung oleh 12% data keseluruhan.
- Dengan *Metode Market Basket Analysis* menggunakan Algoritma *Hash Based* di dapat hasil *item* yang paling banyak dibeli berdasarkan periode Agustus 2022 – Juli 2023 untuk *1-itemset* adalah Sanger Ice (A) Sebesar 42%. Untuk *2-itemset* yaitu Sanger Ice (A) dan Sanger Hot (B) Sebesar 25%. Setelah dilakukan tiga kali *iterasi* tersisa satu alamat pada tabel *hash* dengan *3-itemset*. *Frequent itemset* yang di dapat yaitu Sanger Ice (A), Sanger Hot (B), Sarang Burung (AO) dengan jumlah yang terjual 42 (*Support Count*) atau 12% Dan *Confidence* 49

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Universitas Harapan Medan dan Fakultas Teknik dan Komputer program studi Teknik Informatika atas penelitian yang telah saya lakukan. Terimakasih saya sampaikan kepada orang tua saya, kakak saya, dan dosen pembimbing yang telah membimbing saya sehingga karya ilmiah ini dapat diselesaikan.

DAFTAR PUSTAKA

- [1] J. Wu *et al.*, “An Empirical Study on Customer Segmentation by Purchase Behaviors Using a RFM Model and K -Means Algorithm,” *Math. Probl. Eng.*, vol. 2020, no. November 2017, 2020, doi: 10.1155/2020/8884227.
- [2] E. Umar, D. Manongga, dan A. Iriani, “Market Basket Analysis Menggunakan Association Rule dan Algoritma Apriori Pada Produk Penjualan Mitra Swalayan Salatiga,” *J. Media Inform. Budidarma*, vol. 6, no. 3, hal. 1367, 2022, doi: 10.30865/mib.v6i3.4217.
- [3] R. T. Aldisa, “Penerapan Data Mining Pada Analisa Pola Pembelian Obat Menerapkan Algoritma Hash Based,” *Build. Informatics, Technol. Sci.*, vol. 4, no. 4, hal. 1892–1898, 2023, doi: 10.47065/bits.v4i4.3142.
- [4] R. Fitriana dan U. Chadhiq, “Pengaruh E-Commerce dan Kualitas Produk terhadap Keputusan Pembelian melalui Kepuasan Konsumen sebagai Variabel Intervening,” *AKSES J. Ekon. dan Bisnis*, vol. 17, no. 1, hal. 86–96, 2022, doi: 10.31942/akses.v17i1.6514.
- [5] U. R. Amanda dan D. P. Utomo, “Penerapan Data Mining Algoritma Hash Based Pada Data Pemesanan Buah Impor Cv. Green Uni Fruit,” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 5, no. 1, hal. 86–93, 2021, doi: 10.30865/komik.v5i1.3653.
- [6] F. Panjaitan, A. Surahman, dan T. D. Rosmalasari, “Analisis Market Basket Dengan Algoritma Hash-Based Pada Transaksi Penjualan (Studi Kasus: Tb. Menara),” *J. Teknol. dan Sist. Inf.*, vol. 1, no. 2, hal. 111–119, 2020, doi: 10.33365/jtsi.v1i2.450.
- [7] T. D. Prakoso, I. Ernawati, dan H. B. Seta, “Penemuan Pola Asosiasi Pada Data Restoran Menggunakan Algoritma Hash Based,” *Semin. Nas. Mhs. Ilmu Komput. dan Apl.*, hal. 71–80, 2020.
- [8] A. Rifqy Alfyan, A. Hafidzul Kahfi, M. Rizky Kusumayudha, dan M. Rezki, “Analisis Market Basket Dengan Algoritma Apriori Pada Transaksi Penjualan Di Freshfood,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 4, no. 1, hal. 1–8, 2019.
- [9] E. Dwika Putra, M. Husni Rifqo, dan D. Hardianto, “Apriori Algorithm Implementation on Market Basket Analysis (MBA) of Mobile Phone Accessories Implementasi Algoritma Apriori Pada Market Basket Analysis (MBA) Aksesoris Telepon Seluler,” *J. Kom.*, vol. 2, no. 2, hal. 373–382, 2022, [Daring]. Tersedia pada: <https://doi.org/10.53697/jkomitek.v2i2>
- [10] R. R. Rerung, “Penerapan Data Mining dengan Memanfaatkan Metode Association Rule untuk Promosi Produk,” *J. Teknol. Rekayasa*, vol. 3, no. 1, hal. 89, 2018, doi: 10.31544/jtera.v3.i1.2018.89-98.