

Implementasi Data Mining Penjualan Kosmetik Pada Toko Kazana Menggunakan Metode Algoritma Hast Based

Annisa Fadillah Siregar^{*}, Rizkah Fadillah, Miya Andriyani

Universitas Budi Darma, Medan, Indonesia

Email: ^{1,*}annisaf@univ-bd.ac.id, ²riszkahfadillah52@gmail.com, ³miaandriyani0505@gmail.com

Email Penulis Korespondensi: annisaf@univ-bd.ac.id

Abstrak-Seperti yang sudah diketahui setiap wanita sangat menginginkan yang namanya keindahan atau kecantikan terhadap sesuatu hal apalagi kecantikan pada dirinya sendiri. Pada umumnya semua juga sudah tahu yang dapat membuat wanita cantik secara fisik adalah kosmetik. Penjualan alat-alat kosmetik seperti halnya lipstik, bedak, ataupun alat-alat lainnya. Barang seperti inilah yang sering dibeli wanita pada umumnya. Alat-alat pada kosmetik dapat mengubah penampilan wanita yang biasa pada umumnya menjadi luar biasa. Banyak barang atau alat kosmetik yang dibeli wanita tidak terlepas dari perhitungan data pada setiap penjualan alat-alat kosmetik. Disini kita akan mengetahui cara menghitung penjualan data alat kosmetik. Cara yang tepat untuk mengetahui perhitungan dari alat-alat kosmetik dengan menggunakan Algoritma Hash Based. Pada Algoritma Hash Based ini terdapat nilai minimum support, nilai minimum confidence dan jumlah data. Dengan menggunakan Algoritma Hash Based kita dapat mengetahui perhitungan atau jumlah dari suatu data pada penjualan kosmetik dengan benar. Setelah dilakukan 3 kali iterasi tersisa 1 alamat pada tabel hash dengan 3-itemset. Didapatkan frequent itemset untuk 3-itemset yaitu (Pembersih Wajah, Kutek, Eyeshadow) dengan support count = 2.

Kata Kunci: Data Mining; Penjualan Kosmetik; Algoritma Hash Based

Abstract-As you know, every woman really wants beauty or beauty in something, especially beauty in herself. In general, everyone already knows that what can make a woman physically beautiful is cosmetics. Sales of cosmetic tools such as lipstick, powder, or other tools. Items like this are often purchased by women in general. Tools in cosmetics can change the appearance of an ordinary woman in general to be extraordinary. Many items or cosmetic equipment purchased by women cannot be separated from the calculation of data on each sale of cosmetic equipment. Here we will find out how to calculate sales of cosmetic equipment data. The right way to find out the calculation of cosmetic tools is by using the Hash Based Algorithm. In this Hash Based Algorithm there is a minimum support value, minimum confidence value and amount of data. By using the Hash Based Algorithm we can find out the calculation or amount of data on cosmetic sales correctly. After 3 iterations, 1 address remains in the hash table with 3-itemset. Obtained frequent itemset for 3-itemset, namely (Face Cleanser, Nail Polish, Eyeshadow) with support count = 2.

Keywords: Data Mining; Sales of Cosmetics; Hash Based Algorithm

1. PENDAHULUAN

Kecantikan secara fisik tidak lepas dari yang namanya alat-alat kosmetik, sebab dengan adanya alat kosmetik manusia dapat mengubah fisiknya dari yang biasa saja pada umumnya menjadi yang istimewa. Pembelian alat-alat kosmetik dipengaruhi banyak dikitnya pembeli dapat diketahui dengan harga alat kosmetik. Untuk merk yang relatif murah biasanya dibeli oleh masyarakat yang sederhana atau menengah, sedangkan untuk merk yang relatif mahal biasanya dibeli oleh masyarakat yang memiliki ekonomi tinggi atau tokoh-tokoh publik populer (artis) [1]. Pada dasarnya penjualan alat-alat kosmetik pada toko kosmetik biasanya berhubungan langsung dengan pendataan pada setiap penjualan barang. Pendataan tersebut yang dimaksud dengan menganalisa data suatu barang, menghitung, pencarian data dan pengumpulan data. Hal tersebut didasari dengan adanya pemasukan atau pengeluaran terhadap suatu barang kosmetik tersebut [2].

Saat ini masih terdapat beberapa toko masih menggunakan pendataan yang biasa-biasa saja bisa dikatakan pendataan asal-asalan. Pendataan yang biasa-biasa saja bisa dapat menyebabkan hal yang tidak diinginkan. Salah satu kesalahan pendataan yang tidak diinginkan biasanya kesalahan dalam menghitung pendapatan keuntungan atau kerugian, menganalisa suatu barang, pencarian informasi (alat-alat kosmetik yang sedang populer atau yang sering dipakai saat ini) dalam menjual suatu barang kosmetik. Kesalahan tersebut bisa saja berdampak negatif terhadap toko tersebut sehingga bisa merugikan toko ataupun toko tersebut bangkrut [3]. Saat ini banyak diketahui beberapa teknik-teknik dalam pendataan suatu barang dengan benar. Teknik ini yang berguna untuk mengelolah data dalam jumlah yang sangat besar yang bisa kita kenal dengan nama Data Mining. Data mining digunakan untuk menggali suatu pengetahuan informasi dengan cara menganalisa dalam jumlah yang sangat besar. Dengan menggunakan teknik ini, dapat mengkonversi data-data yang besar menjadi data yang bernilai yang berisi suatu informasi yang berguna. Pentingnya suatu aturan asosiasi dapat diketahui dengan dua parameter yaitu *support* dan *confidence*[4].

Salah satu aturan yang dapat digunakan pada pendataan suatu barang yaitu aturan asosiasi. Aturan ini tidak melibatkan prediksi secara langsung. Penggunaan aturan asosiasi dapat digunakan untuk menemukan keterkaitan atribut, produk atau peristiwa. Salah satu tahap asosiasi yang menarik yaitu analisis pola frekuensi tinggi. Analisa pola frekuensi tinggi adalah untuk mencari itemset yang memenuhi syarat minimum dari nilai *support count* dalam database. Untuk memenuhi itemset nilai minimum *support* dari kandidat itemset dalam jumlah yang besar dapat mendominasi performa data mining secara keseluruhan[5]. Algoritma yang digunakan dalam penelitian ini yaitu Algoritma Hash Based. Algoritma Hash Based dapat menjadi solusi untuk menentukan frekuensi itemset dari kandidat itemset secara optimal. Algoritma ini dapat mengurangi jumlah kandidat k-itemset pada awal. Sehingga pendataan dari suatu jumlah

barang dapat tersusun dengan baik [6]. Pada tahap ini penelusuran pada data untuk melakukan perhitungan *support count* dari setiap kandidat *k-itemset* dan pada saat yang sama teknik ini mencari informasi *k-itemset* dengan menggunakan *hash* pada semua data yang diproses terhadap semua hasil kemungkinan yang disimpan pada tabel *hash*. Informasi yang telah disimpan pada tabel *hash* akan dimasukkan kedalam *bucket-bucket*, Setiap *bucket* akan menyimpan *support count* pada *bucket* tersebut, Setelah itu akan dieliminasi untuk *bucket* yang tidak memenuhi nilai *support* minimum [7].

Berdasarkan penelitian terkait yang dilakukan oleh Feresia Panjaitan dkk tahun 2020 meneliti mengenai penerapan algoritma hash-based pada analisis market basket transaksi penjualan di TB. Menara. Penelitian menghasilkan 2 Frequent itemset serta association rules dari data penjualan bahan bangunan[6]. Penelitian yang dilakukan oleh Triaji Doni Prakoso dkk tahun 2020 membahas mengenai penerapan algoritma hash-based dalam menemukan pola asosiasi pada data restoran. Penelitian menyimpulkan hasil yang diperoleh yaitu rekomendasi itemset yang diperoleh dari data transaksi[8].

Berdasarkan penjelasan yang telah dijabarkan penulis mengangkat topik yang berjudul “Implementasi Data Mining Penjualan Kosmetik Pada Toko Kazana Menggunakan Metode Algoritma Hash Based”.

2. METODOLOGI PENELITIAN

2.1 Penjualan

Pada proses pembelian, berkaitan langsung dengan pencarian informasi dan nilai-nilai dari sumber untuk memenuhi kebutuhan dan kegiatan yang dirasakan. Semakin banyaknya informasi yang diperoleh maka semakin banyak pula pengetahuan bagi konsumen dalam hal memilih suatu produk untuk menentukan kualitas dan harga dari suatu barang. Penjualan dalam hal ini banyak kesalahpahaman seseorang menanggapi apa itu arti dari penjualan yaitu penjualan dan pemasaran, padahal seharusnya penjualan dapat iartikan sebagai pemasaran dari suatu perusahaan tersebut[9]. Setelah mendapatkan informasi yang berasal dari kegiatan yang dilakukan di pasar, maka diharapkan adanya suatu penjualan terhadap barang yang di pesan. Dalam hal ini kegiatan penjualan yang dilakukan antara penjual dan pembeli yang mempengaruhi beberapa faktor, antara lain : Kondisi dan kemampuan penjual, Kondisi pasar, Modal dan kondisi organisasi lainnya[10].

2.2 Algoritma Hash Based

Algoritma Hash Based merupakan teknik *hashing* untuk menyaring keluar *itemset* yang tidak penting untuk pembangkitan *itemset* selanjutnya. Ketika *super count* kandidat *k-itemset* dihitung menggunakan database, Algoritma Hash Based mengumpulkan informasi mengenai $(k+1)$ -*itemset* dengan cara seluruh kemungkinan $(k+1)$ -*itemset* kedalam hash table dengan menggunakan fungsi hash. Algoritma Hash Based juga dapat menjadi solusi menentukan kandidat itemset secara optimal. Berdasarkan setiap *bucket* pada hash table berisi angka, akan dibangun bit vektor yang bernilai 1 pada angka *bucket* yang bersangkutan lebih besar atau sama dengan minimum *support*. Pada bagian pembangkitan kandidat, setelah menghitung $C_k = L_{k-1} * L_{k-1}$, setiap *k-itemset* diperiksa apakah *itemset* tersebut sama dengan satu. Apabila *k-itemset* tidak sama dengan satu maka *itemset* tersebut tidak dapat digunakan. Penggunaan Hash table ini dapat mengurangi jumlah pada kandidat *k-itemset* sehingga mampu mengurangi nilai komputasi dari pembangkitan *itemset*. Menurut [11] Ada tiga bagian utama untuk dapat membangkitkan *frequent itemset algoritma hash based*, yaitu [12]:

1. Algoritma Hash Based akan menghasilkan C_1 (kandidat 1-*itemset*) dan L_1 (Large 1-*itemset*) dari *database*. Untuk menghitung *support count* dari *itemset* ini ditelusuri seluruh transaksi untuk kandidat 1-*itemset*. Pada tahap ini *hash tree* untuk C_1 dibangun untuk menghitung efisiensi *support count*. Algoritma *hash based* memeriksa apakah setiap *item* sudah ada dalam *hash tree*. Algoritma juga akan membangun *hash table* untuk 2 *itemset* yang bertujuan untuk mengurangi banyaknya kandidat 2-*itemset*, C_2 .
2. Kumpulan kandidat *itemset* C_k dibangkitkan pada *hash table* yang dibuat pada iterasi sebelumnya. Setelah itu ditentukan *frequent itemset* untuk mengurangi ukuran *database* pada pembangkitan *itemset* selanjutnya. Ada dua fase yang terbagi untuk kandidat $(k+1)$ -*itemset*, yaitu :
 - a. Fase Pertama untuk membangkitkan kandidat K -*itemset* C_k berdasarkan pada Hash Base H_k , Proses ini terjadi pada prosedur *gen-candidate*
 - b. Fase Kedua Algoritma akan menghitung *support* pada kandidat *itemset* untuk mengurangi ukuran pada setiap transaksi, proses ini terjadi pada *super count*.
3. Algoritma *Hash Based* diganti dengan Algoritma Apriori karena Algoritma *Hash Based* tidak efisien dibanding dengan Algoritma Apriori. Berikut ini adalah rumus hash based pada table level kedua dan ketiga. Rumus Hash pada level kedua $H\{XY\} = ((\text{Order of } X) * 10 + (\text{Order of } Y) \text{ mod bil prima}$. Rumus hash pada level ketiga $H\{XYZ\} ((\text{order of } X * 100) + (\text{order of } Y * 10) + \text{order of } Z) \text{ mod bil prima}$.

Keterangan :

Order of X = Perwakilan nilai X

Order of Y = Perwakilan nilai Y

Order of Z = Perwakilan nilai Z

2.3 Tahapan Penelitian

Pada tahapan penelitian ini, penulis dapat melakukan mengumpulkan data-data. Proses ini dilakukan dengan pengumpulan data dengan cara melakukan studi lapangan atau studi pustaka. Pengumpulan data-data yang telah diperoleh akan dianalisa agar mengetahui kebutuhan dari proses asosiasi. Adapun beberapa penjelasan singkat pada tahap penelitian sebagai berikut.

1. Mengidentifikasi Masalah

Hal pertama yang harus dilakukan yaitu Mengidentifikasi Masalah. Pada tahap ini yang akan dilakukan adalah dengan mencari, mengumpulkan, Mencatat data dan informasi dari hasil yang dilakukan dilapangan.

2. Menganalisa Masalah

Tahap selanjutnya yaitu Menganalisa Masalah. Pada tahap ini akan dilakukan pemodelan dan pembersihan data dengan tujuan untuk mencari informasi penting dari data-data yang dikumpulkan, Sehingga mendapatkan kesimpulan dari pengambilan data.

3. Mengumpulkan Data

Dalam pengumpulan informasi dapat dilakukan observasi data dengan cara pengamatan langsung terhadap toko kosmetik tersebut sehingga kasus yang didapat bisa dikenal secara jelas. Selanjutnya mencoba untuk menginterview dengan tujuan agar bisa mendapatkan informasi atau data yang diperlukan. Studi pustaka dapat juga dilakukan dengan membaca buku-buku yang berhubungan dengan kasus dengan tujuan untuk mencari data informasi terhadap data-data yang didapat. Data yang berhasil didapat, selanjutnya dicoba untuk menggunakan Algoritma Hash-Based.

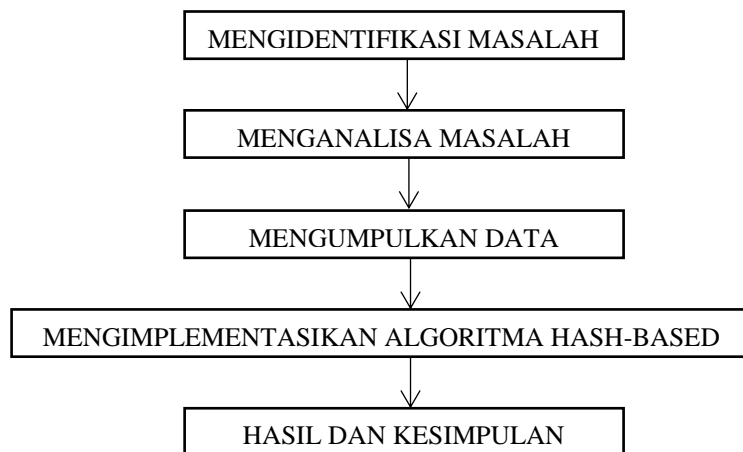
4. Mengimplementasikan Algoritma Hash-Based

Pada tahap ini dilakukan menerapkan suatu model atau metode dari sistem dengan menggunakan Algoritma Hash-Based. Dengan tujuan agar diperoleh nilai *support* minimum dan nilai *confidence* dengan kandidat $(k+1)$ -itemset yang berasal dari *large k-itemset*.

5. Hasil dan Kesimpulan

Pada tahap ini dilakukan pegujian dan perancangan sistem untuk mendapatkan hasil dan kesimpulan. Sistem diuji dengan pemodelan dan prosedur-proseur yang nantinya dilakukan eksplorasi dari data-data, sehingga mendapatkan data yang tersembunyi.

Berikut gambar kerangka penelitian dari penjelasan di atas:



Gambar 1. Kerangka Penelitian

3. HASIL DAN PEMBAHASAN

3.1 Analisa Masaalah

Analisa masalah merupakan tahap awal yang dilakukan untuk memecahkan permasalahan yang sedang terjadi. Pada tahap analisa ini sangat penting karena proses analisis yang akurat akan menghasilkan sebuah system yang dapat digunakan untuk pengguna. Dalam hal ini permasalahan yang akan dibahas adalah menganalisa pendataan barang kosmetik pada Toko Kosmetik Kazana dalam mengelola pendataan barang kosmetik dengan tujuan Membantu Toko Kazana dalam menganalisa barang kosmetik agar tidak terjadi kesalahan dalam penjualan barang serta mengatasi masalah atau kesalahan dalam pendataan agar tidak mengalami kerugian.

3.2 Penerapan Algoritma Hash-Based

Pada Data transaksi yang sudah dikumpulkan dari toko kosmetik kazana dengan mengambil 10 sampel alat kosmetik.

Tabel 1. Data Transaksi

ID Transaksi	Itemset
T1	Pembersih Wajah, Kutek
T2	Lipstik, Bedak, Kutek
T3	Pemebersih Wajah, Lipstik, Kutek, Eyeshadow
T4	Blush On, Lipstik, Kutek
T5	Pembersih Wajah, Eyeshadow
T6	Lipstik
T7	Eyelinier, Sunblock
T8	Pembersih Wajah, Maskara, Kutek, Eyeshadow
T9	Bedak, Blush On, Eyelinier
T10	Kutek, Lipstik

Selanjutnya akan dilakukan proses hashing terhadap kandidat 1-itemset untuk memasukan setiap itemset kedalam bucket pada tabel hash. Pada proses hashing ini menggunakan rumus sebagai berikut.

$$h(x) = (\text{order of item } x) \bmod n \tag{1}$$

keterangan :

h = alamat bucket pada tabel hash

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

Tabel 2. Order of Item

Item	Order
Pembersih Wajah	1
Kutek	2
Lipstik	3
Bedak	4
Blush On	5
Eyeshadow	6
Eyelinier	7
Sunblock	8
Maskara	9

Mencari Alamat pada tabel hash untuk 1-itemset

$$h(\text{Pembersih Wajah}) = (1) \bmod 19 = 1$$

$$h(\text{Kutek}) = (2) \bmod 19 = 2$$

$$h(\text{Lipstik}) = (3) \bmod 19 = 3$$

$$h(\text{Bedak}) = (4) \bmod 19 = 4$$

$$h(\text{Blush On}) = (5) \bmod 19 = 5$$

$$h(\text{Eyeshadow}) = (6) \bmod 19 = 6$$

$$h(\text{Eyelinier}) = (7) \bmod 19 = 7$$

$$h(\text{Sunblock}) = (8) \bmod 19 = 8$$

$$h(\text{Maskara}) = (9) \bmod 19 = 9$$

Tabel 3. Tabel Hash 1-itemset

Address	Itemset	Link	Support Count
1	Pembersih Wajah	T1, T3, T5, T8	4
2	Kutek	T1, T2, T3, T4, T8, T10	6
3	Lipstik	T2, T3, T4, T6, T10	5
4	Bedak	T2, T9	2
5	Blush On	T4, T9	2
6	Eyeshadow	T5, T6, T8	3
7	Eyelinier	T7, T9	2
8	Sunblock	T7	1
9	Maskara	T8	1

Setelah dilakukan perhitungan hash maka akan dihasilkan data pada Tabel 2. Itemset menempati alamat hash sesuai dengan hasil setelah dilakukan perhitungan menggunakan rumus *hashing*. Alamat yang akan ditempati oleh itemset akan menjadi *node* dan membuat *link* yang mengarah kepada transaksi-transaksi yang telah mengandung itemset secara berurutan sehingga akan terbentuk *link list*. Dengan menentukan minimum *support count* = 2 dihasilkan *frequent* 1-itemset seperti pada Tabel .

Tabel 4. Frequent 1-itemset

Item Set	Support Count
Pembersih Wajah	4
Kutek	6
Lipstik	5
Bedak	2
Blush On	2
Eyeshadow	3
Eyelinier	2

Langkah Selanjutnya adalah mencari *Frequent 2-itemset*. Sebelum mencari *Frequent 2-itemset* terlebih dahulu ditentukan semua kemungkinan 2-itemset yang terbentuk dari *Frequent 1-itemset*. Kandidat 2-itemset yang terbentuk dari *frequent 1-itemset* yaitu, (Pembersih Wajah, Kutek), (Pembersih Wajah, Lipstik), (Pembersih Wajah, Bedak), (Pembersih Wajah, *Blush On*), (Pembersih Wajah, *Eyeshadow*), (Pembersih Wajah, *Eyelinier*), (Kutek, Lipstik), (Kutek, Bedak), (Kutek, *Blush On*), (Kutek, *Eyeshadow*), (Kutek *Eyelinier*), (Lipstik, Bedak), (Lipstik, *Blush On*), (Lipstik, *Eyeshadow*), (Lipstik, *Eyelinier*), (Bedak, *Blush On*), (Bedak, *Eyeshadow*), (Bedak, *Eyelinier*), (*Blush On*, *Eyeshadow*), (*Blush On*, *Eyelinier*), (*Eyeshadow*, *Eyelinier*). Selanjutnya memasukan 2-itemset kedalam tabel hash dengan menggunakan rumus hash sebagai berikut :

$$h(k) = ((\text{order of } x) * 10 + \text{order of } y) \bmod n \quad (2)$$

Mencari Alamat pada tabel hash untuk 2-itemset

$h(\text{Pembersih Wajah, Kutek})$	$= ((1) * 10 + 2) \bmod 19 = 12$
$h(\text{Pembersih Wajah, Lipstik})$	$= ((1) * 10 + 3) \bmod 19 = 13$
$h(\text{Pembersih Wajah, Bedak})$	$= ((1) * 10 + 4) \bmod 19 = 14$
$h(\text{Pembersih Wajah, Blush On})$	$= ((1) * 10 + 5) \bmod 19 = 15^*$
$h(\text{Pembersih Wajah, Eyeshadow})$	$= ((1) * 10 + 6) \bmod 19 = 16^*$
$h(\text{Pembersih Wajah, Eyelinier})$	$= ((1) * 10 + 7) \bmod 19 = 17$
$h(\text{Kutek, Lipstik})$	$= ((2) * 10 + 3) \bmod 19 = 4$
$h(\text{Kutek, Bedak})$	$= ((2) * 10 + 4) \bmod 19 = 5$
$h(\text{Kutek, Blush On})$	$= ((2) * 10 + 5) \bmod 19 = 6$
$h(\text{Kutek, Eyeshadow})$	$= ((2) * 10 + 6) \bmod 19 = 7^*$
$h(\text{Kutek Eyelinier})$	$= ((2) * 10 + 7) \bmod 19 = 8$
$h(\text{Lipstik, Bedak})$	$= ((3) * 10 + 4) \bmod 19 = 15^*$
$h(\text{Lipstik, Blush On})$	$= ((3) * 10 + 5) \bmod 19 = 16^*$
$h(\text{Lipstik, Eyeshadow})$	$= ((3) * 10 + 6) \bmod 19 = 17$
$h(\text{Lipstik, Eyelinier})$	$= ((3) * 10 + 7) \bmod 19 = 18^*$
$h(\text{Bedak, Blush On})$	$= ((4) * 10 + 5) \bmod 19 = 7^*$
$h(\text{Bedak, Eyeshadow})$	$= ((4) * 10 + 6) \bmod 19 = 8$
$h(\text{Bedak, Eyelinier})$	$= ((4) * 10 + 7) \bmod 19 = 9$
$h(\text{Blush On, Eyeshadow})$	$= ((5) * 10 + 6) \bmod 19 = 18^*$
$h(\text{Blush On, Eyelinier})$	$= ((5) * 10 + 7) \bmod 19 = 19$
$h(\text{Eyeshadow, Eyelinier})$	$= ((6) * 10 + 7) \bmod 19 = 10$

Pada perhitungan hashing ditemukan adanya *collison* (lebih dari satu itemset yang memiliki alamat hash yang sama). Pada kasus ini *collison* terjadi pada alamat ke-7, alamat ke-15, alamat ke-16 dan alamat ke-18. Ketika *collison* terjadi maka harus dilakukan pengecekan untuk mencari alamat *bucket* yang masih kosong. Setelah dilakukan pengecekan dan ditemukan indikasi bahwa tabel hash sudah separuhnya terisi maka perlu dilakukan rehashing ulang dengan banyak alamat 2 kali dari banyak alamat sebelumnya. Untuk itu dilakukan penyelesaian masalah dengan menggunakan rumus yang berbeda.

$$h(k) = ((\text{order of } x) * 10 + \text{order of } y) \bmod j \quad (3)$$

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

$h(\text{Pembersih Wajah, Kutek})$	$= ((1) * 10 + 2) \bmod 39 = 12$
$h(\text{Pembersih Wajah, Lipstik})$	$= ((1) * 10 + 3) \bmod 39 = 13$
$h(\text{Pembersih Wajah, Bedak})$	$= ((1) * 10 + 4) \bmod 39 = 14$
$h(\text{Pembersih Wajah, Blush On})$	$= ((1) * 10 + 5) \bmod 39 = 15$
$h(\text{Pembersih Wajah, Eyeshadow})$	$= ((1) * 10 + 6) \bmod 39 = 16$
$h(\text{Pembersih Wajah, Eyelinier})$	$= ((1) * 10 + 7) \bmod 39 = 17$
$h(\text{Kutek, Lipstik})$	$= ((2) * 10 + 3) \bmod 39 = 23$
$h(\text{Kutek, Bedak})$	$= ((2) * 10 + 4) \bmod 39 = 24$
$h(\text{Kutek, Blush On})$	$= ((2) * 10 + 5) \bmod 39 = 25$

h(Kutek, Eyeshadow)	= ((2)*10+6)mod 39 =26
h(Kutek Eyeliner)	= ((2)*10+7)mod 39 =27
h(Lipstik, Bedak)	= ((3)*10+4)mod 39 =34
h(Lipstik, Blush On)	= ((3)*10+5)mod 39 = 35
h(Lipstik, Eyeshadow)	= ((3)*10+6)mod 39 =36
h(Lipstik, Eyeliner)	= ((3)*10+7)mod 39 =37
h(Bedak, Blush On)	= ((4)*10+5)mod 39 =7
h(Bedak, Eyeshadow)	= ((4)*10+6)mod 39 =8
h(Bedak, Eyeliner)	= ((4)*10+7)mod 39 =9
h(Blush On, Eyeshadow)	= ((5)*10+6)mod 39 =18
h(Blush On, Eyeliner)	= ((5)*10+7)mod 39 =19
h(Eyeshadow, Eyeliner).	= ((6)*10+7)mod 39 =10

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

Tabel 5. Tabel Hash 2 – Itemset

Address	Itemset	Link	Support Count
0	-	-	-
...			
7	Bedak, Blush On	T9	1
8	Bedak, Eyeshadow	-	0
9	Bedak, Eyeliner	T9	1
10	Eyeshadow, Eyeliner	-	0
...			
12	Pembersih Wajah, Kutek	T1, T3, T8	3
13	Pembersih Wajah, Lipstik	T3	1
14	Pembersih Wajah, Bedak	-	0
15	Pembersih Wajah, Blush On	-	0
16	Pembersih Wajah, Eyeshadow	T5, T8	2
17	Pembersih Wajah, Eyeliner	-	0
18	Blush On, Eyeshadow	-	0
19	Blush On, Eyeliner	T9	1
...			
23	Kutek, Lipstik	T2, T3, T4, T10	4
24	Kutek, Bedak	T2	1
25	Kutek, Blush On	T4	1
26	Kutek, Eyeshadow	-	0
27	Kutek, Eyeliner	-	0
...			
34	Lipstik, Bedak	T2	1
35	Lipstik, Blush On	T4	1
36	Lipstik, Eyeshadow	T6	1
37	Lipstik, Eyeliner	-	-
...			
39	-	-	-

Dari table hash tersebut akan dilakukan eliminasi itemset terhadap alamat bucket dengan *support count* <2 untuk mendapatkan *frequent 2* – itemset.

Tabel 6. Frequent 2 – itemset

Itemset	Support Count
Pembersih Wajah, Kutek	3
Pembersih Wajah, Eyeshadow	2
Kutek, Lipstik	4

Setelah mendapatkan *frequent 2*-itemset, Maka langkah selanjutnya adalah mencari *frequent 3*-itemset yang terbentuk. 3-itemset yang terbentuk yaitu : (Pembersih Wajah, Kutek, *Eyeshadow*), (Pembersih Wajah, Kutek, Lipstik), (Pembersih Wajah, *Eyeshadow*, Lipstik). Untuk memasukan 3-itemset tersebut kedalam tabel Hash maka digunakan rumus sebagai berikut :

$$H(k) = ((\text{order of } X) * 100 + (\text{order of } Z)) \bmod j \tag{4}$$

mencari alamat pada tabel hash untuk 3-itemset:

$$H(\text{Pembersih Wajah, Kutek, Eyeshadow}) = ((1)*100+(2)*10+3) \bmod 39 = 6$$

$$H(\text{Pembersih Wajah, Kutek, Lipstik}) = ((1)*100+(2)*10+4) \bmod 39 = 7$$

$$H(\text{Pembersih Wajah, Eyeshadow, Lipstik}) = ((1)*100+(2)*10+5) \bmod 39 = 8$$

Tabel 7. Tabel Hash 3-Itemset

Address	Itemset	Link	Support Count
0			
...			
6	Pembersih Wajah, Kutek, Eyeshadow	T3, T8	2
7	Pembersih Wajah, Kutek, Lipstik	T3	1
8	Pembersih Wajah, Eyeshadow, Lipstik	T3	1

Tabel 8. Frequent 3-itemset

Itemset	Super Count
Pembersih Wajah, Kutek, Eyeshadow	2

Setelah dilakukan 3 kali iterasi tersisa 1 alamat pada tabel hash dengan 3-itemset. Didapatkan frequent itemset untuk 3-itemset yaitu (Pembersih Wajah, Kutek, Eyeshadow) dengan support count = 2.

4. KESIMPULAN

Dari penelitian yang sudah dilakukan didapatkan kesimpulan yaitu dalam menentukan frequent itemset tahap yang paling berpengaruh terhadap performa data mining, terutama jika kandidat itemset berjumlah sangat besar. Hanya dilakukan 1 scan database pada Algoritma *Hash Based*. Untuk Selanjutnya, itemset tersebut akan dimasukkan kedalam *bucket* pada tabel hash menggunakan operasi hash dengan ketentuan setiap alamat *bucket* diisi oleh 1 itemset. Dengan begitu, performa data mining meningkat karena scan database dilakukan hanya sekali pada awal iterasi sehingga tidak lagi dilakukan scan database pada setiap mencari *frequent* k-itemset. Pada setiap iterasi yang tidak memenuhi *support* minimum *count* akan dilakukan eliminasi. Dari 10 sampel data didapatkan 2 itemset yang memenuhi *support* minimum *count* setelah dilakukan iterasi hingga 3-itemset. Setelah dilakukan 3 kali iterasi tersisa 1 alamat pada tabel hash dengan 3-itemset. Didapatkan frequent itemset untuk 3-itemset yaitu (Pembersih Wajah, Kutek, Eyeshadow) dengan support count = 2.

REFERENCES

- [1] N. Hidayah, "Analisis Strategi Digital Marketing Dalam Membantu Penjualan Living Space Dan Efo Store," 2018.
- [2] M. Ulkhaiir, "Implementasi Data Mining Penjualan Kosmetik Dengan Algoritma Apriori." Prodi Teknik Informatika, 2020.
- [3] D. Irawan and K. D. Perindustrian, "Industri Produk Tekstil (Apd) Jawa Timur Meningkatkan Di Tengah Pandemi Covid-19," *Merdeka Berpikir Catatan Hari. Pandemi Covid*, vol. 19, p. 111, 2020.
- [4] 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," *KLIK - Kumpul. J. ILMU Komput.*, 2018, doi: 10.20527/klik.v5i1.100.
- [5] A. H. SIREGAR, "Analisis Prediksi Aturan Asosiasi Menggunakan Algoritma Ct-Pro Dan Algoritma Hash-Based Dalam Kasus Kekerasan Pada Anak," 2020.
- [6] F. Panjaitan, A. Surahman, and 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, pp. 111–119, 2020.
- [7] M. M. Rokhimin, M. T. A. C. Widiyanto, and H. Sikumbang, "Implementasi Metode Max-Miner dan Market Basket Analisis untuk Rekomendasi Produk di PT. Astra Otopart." INSTITUT TEKNOLOGI PLN, 2020.
- [8] T. D. Prakoso *et al.*, "PENEMUAN POLA ASOSIASI PADA DATA RESTORAN MENGGUNAKAN ALGORITMA HASH BASED," pp. 71–80, 2020.
- [9] A. A. Tri Susilo, L. Sunardi, and Y. Waruwu, "PENERAPAN ALGORITMA APRIORI PADA DATA PENJUALAN KOSMETIK DI TOKO SHARLY KOTA LUBUKLINGGAU," *JUTIM (Jurnal Tek. Inform. Musirawas)*, 2019, doi: 10.32767/jutim.v4i2.648.
- [10] C. Adiwihardja, N. Hardi, and W. Widyastuty, "Implementasi Data Mining Penjualan Kosmetik Pada Toko Zahrani Menggunakan Algoritma Apriori," *J. Speed – Sentra Penelit. Eng. dan Edukasi*, 2019.
- [11] P. R. Putri, "MARKET BASKET ANALYSIS DENGAN MENGGUNAKAN ALGORITMA HASH-BASED PADA TRANSAKSI PENJUALAN APOTEK UNTUK MENERAPKAN KONSEP CROSS-SELLING," *Biology (Basel)*, 2016.
- [12] A. J. P. Sibarani, "Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 2, pp. 262–276, 2020, doi: 10.35957/jatisi.v7i2.195.