

Penerapan Data Mining Algoritma Hash Based Pada Data Pemesanan Buah Impor Cv. Green Uni Fruit

Ulva Rizky Amanda¹, Dito Putro Utomo¹

¹Fakultas Ilmu Komputer Dan Teknologi Infomasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia
Email: ¹ulvarizkyamanda@gmail.com, ²ditoputro12@gmail.com
^{*)} Email Penulis Korespondensi

Abstrak– Data mining adalah proses yang menggunakan statistika, kecerdasan buatan, dan *mechine learning* untuk mengekstraksi dan mengidentifikasi data menjadi bermanfaat. CV. Green Uni Fruit merupakan perusahaan distributor dan importir buah-buahan memiliki pola pemesanan dengan data yang terarsip namun tidak dimanfaatkan dan hanya memenuhi ruang penyimpanan. Fungsi data mining asosiatif dapat digunakan untuk menemukan kombinasi atau aturan asosiatif suatu data sehingga data yang sebelumnya hanya terarsip, dapat ditemukan informasi dan hubungan dengan variabel lainnya. Salah satu teknik yang dapat digunakan pada proses mining rule asosiatif adalah algoritma Hash Based. Pada proses ini akan dilakukan pembangkitan kandidat rule menggunakan table hash based dan proses pruning pada data transaksi. Penggunaan algoritma hash based pada proses mining rule asosiatif menjadikan penggunaan data penyimpanan lebih efektif dan efisien.

Kata Kunci: Algoritma, Data Mining, Hash Based, Asosiatif, Pemesanan

Abstract– *Data mining is a process that uses statistics, artificial intelligence, and machine learning to extract and identify useful. CV. Green Uni Fruit is a fruit distributor and importer company that has an ordering pattern with archived data but is not utilized and only fills storage space. Associative data mining functions can be used to find combinations or associative rules of data so that previously only archived data, information and relationships with other variables can be found. One technique that can be used in the associative rule mining process is the Hash Based algorithm. In this process, candidate rule will be generated using a hash based table and pruning process on transaction data. The use of hash algorithms based on the associative rule mining process makes the use of data storage more effective and efficient.*

Keywords: *Algorithm, Data Mining, Hash Based, Asosiatif, Ordering*

1. PENDAHULUAN

Kebutuhan informasi yang akurat dan cepat di zaman teknologi sekarang ini sangat diperlukan untuk mendukung operasional pemesanan berbagai bisnis dengan efektif dan efisien. Proses pemesanan dalam bisnis memiliki siklus tersendiri di mana saat ada pihak pemesanan maka terlebih dahulu bagian admin harus membuat surat pemesanan *purchasing order*. Setelah membuat daftar barang yang dibeli, perusahaan melakukan pemesanan dengan mengirimkan surat pesanan kepada *supplier*. Perusahaan sebaiknya tidak mengirimkan surat pemesanan secara mendadak untuk menghindari adanya selisih dalam proses muat barang dan pendistribusian barang nanti. Oleh karena itu harus admin harus bisa mengestimasi waktu pengiriman ketika melakukan proses pemesanan tersebut.

CV. Green Uni Fruit merupakan salah satu perusahaan berkembang yang berdomisili di kota Medan. Perusahaan ini bergerak sebagai distributor dan importir buah – buahan dari berbagai Negara. Green Uni Fruit yang didirikan oleh Bapak Suwandi Cong selaku *owner*. Beliau yang dibantu oleh rekan-rekannya, akhirnya bekerja sama dalam mengolah perusahaan ini dengan tujuan demi menyalurkan dan memenuhi kebutuhan gizi masyarakat maka mereka membangun perusahaan.

CV. Green Uni Fruit dalam menjalankan aktivitas operasionalnya pendataan pemesanan buah oleh customer disimpan kedalam database. Data-data tersebut makin lama banyak dan hanya terarsip saja. Data yang tersimpan tersebut hanya menghabiskan ruang penyimpanan dan tidak memberikan manfaat pada perusahaan. Oleh karena itu, dibutuhkan suatu teknik dalam penggalian informasi / pengetahuan dari data tersebut.

Perkembangan teknologi informatika saat ini telah mencakup ke berbagai bidang kehidupan. Salah satu bentuk teknologi informatika telah masuk dalam kehidupan seperti penggunaan data mining dalam sistem pembelian dan penjualan pada suatu perusahaan. Dalam data mining terdapat beberapa metode sehingga himpunan data tersebut dapat memiliki pengetahuan yaitu metode prediksi, estimasi, klasifikasi, klastering, dan asosiasi. Dalam metode data mining asosiasi terdapat salah satu algoritma dalam bentuk *hased based*. Algoritma *hased based* sendiri merupakan pengembangan algoritma apriori yang sudah ada sebelumnya. Pengembangan dilakukan pada proses pembangkitan kandidat rule menggunakan *tabel hash* dan proses *pruning* pada data transaksi. Dengan menggunakan algoritma *hased based* proses *mining rule asosiasi* menjadi lebih cepat dan efisien dalam penggunaan memori. Manfaat yang dimiliki *hased based* menjadikan proses mining rule asosiasi yang lebih cepat dan penggunaan memori yang lebih efisien, penulis melihat potensi dari penggunaan algoritma *hased based* untuk menjadi solusi dalam permasalahan yang ada pada CV. Green Uni Fruit.

Penelitian sebelumnya yang dilakukan oleh Farha Ramadhan di tahun 2017 dengan judul Implementasi Algoritma *Hash Based* Terhadap Aturan *Asosiasi* Untuk Menentukan *Frequent Itemset* Studi Kasus Rumah Makan Seafood “Kita” mendapatkan hasil bahwa Algoritma *hash based* menjadi solusi bagi kelemahan algoritma apriori didalam menentukan *frequent itemset*[1].

Penelitian lainnya yang dilakukan oleh Feresia Panjaitan, dkk ditahun 2020 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[2].

Penelitian yang dilakukan oleh Triaji Doni Prakoso,dkk tahun 2020 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 item set[3].

Penelitian lainnya yang telah dilakukan oleh Orisky Sitra Arifah Destiyati dan Eko Aribowo pada tahun 2015 dengan judul penelitian Analisis Perbandingan Algoritma Apriori Dan Algoritma *Hash Based* Pada *Market Basket Analysis* Di Apotek UAD hasil yang didapatkan algoritma *hash based* terjadi *collision*, maka waktu yang dibutuhkan dalam menambang data lebih lama dibandingkan waktu yang dibutuhkan algoritma apriori. Selain itu setelah langkah menempatkan item pada hash tabel sesuai dengan alamatnya, langkah yang dilakukan *hash based* sama dengan langkah yang dilakukan apriori[4].

Penelitian yang lainnya dilakukan oleh R D Yulanda,dkk pada tahun 2019 dengan judul *Association Rules With Apriori Algorithm And Hash-Based Algorithm* didapatkan hasil bahwa algoritma *hash based* lebih efisien karena mereduksi pasangan *itemset* yang lebih baik daripada algoritma apriori[5].

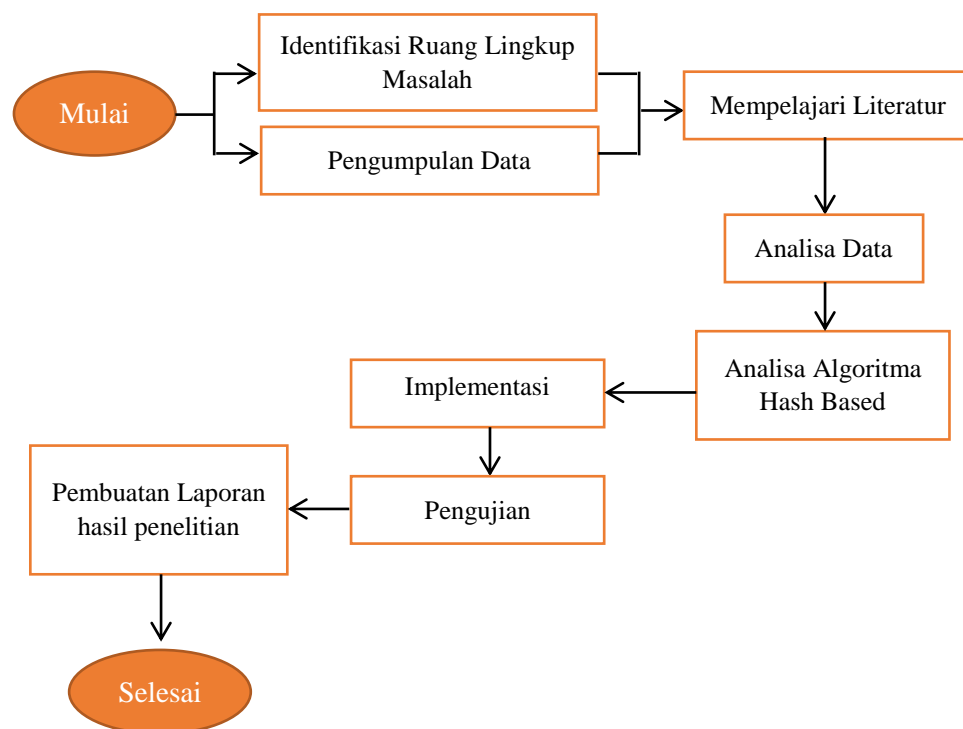
2. METODE PENELITIAN

2.1 Tahapan Penelitian

Metodologi penelitian merupakan sekumpulan kegiatan, prosedur atau tahapan yang dilakukan dalam melakukan sebuah penelitian. Metodologi penelitian dibuat agar penelitian ini lebih terkonsep dan terarah sehingga sesuai dengan tujuan penelitian. Dalam tahap ini akan dijelaskan mengenai tahapan-tahapan penelitian yang harus dilakukan agar mempermudah dalam penyelesaian terhadap permasalahan yang ada, serta melakukan pengumpulan data yang berkaitan dengan penelitian untuk memperoleh data-data dan informasi yang dibutuhkan. Dalam mengumpulkan data dan informasi yang dibutuhkan penulis melakukan penelitian secara langsung pada CV. Green Fruit dengan melakukan observasi dan wawancara dengan pihak perusahaan yang berkaitan dengan data pemesanan buah.

Dalam kegiatan observasi penulis melakukan wawancara dengan mempertanyakan permasalahan yang terjadi dalam penelitian, yaitu tentang bagaimana pemesanan data buah pada CV. Green Fruit. Setelah melakukan wawancara penulis mendapatkan data dan informasi berupa data pemesanan buah.

Setelah mendapatkan informasi dan data yang dibutuhkan, data-data yang diperoleh nantinya akan digunakan dalam menyelesaikan penelitian ini. Dibantu dengan penerapan data mining dengan algoritma *hash based*. Di bawah ini merupakan alur dari tahapan penelitian serta proses pengumpulan data yang dilakukan, adalah sebagai berikut:



Gambar 1. Tahapan Penelitian

2.2 Data Mining

Data mining adalah proses yang menggunakan *statistic*, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi yang bermanfaat. Data mining didefinisikan sebagai proses penemuan pola dalam data[6]. Berdasarkan tugasnya, data mining dikelompokkan menjadi deskripsi, estimasi prediksi, klasifikasi, klastering, dan

asosiasi. Suatu proses penambahan informasi penting dari suatu data. Informasi penting ini didapat dari suatu proses yang amat rumit seperti menggunakan *artificial intelligence*, teknik statistik, ilmu matematika, *machine learning* dan lain sebagainya.

Jadi, berdasarkan penjelasan diatas dapat disimpulkan bahwa data mining merupakan suatu proses penambahan data dalam jumlah data yang sangat besar dengan menggunakan metode statistika, matematika, hingga memanfaatkan teknologi *artificial intelligence* terkini. Menurut para ahli tujuan dari penambahan data ini untuk mengekstraksi serta mengidentifikasi suatu data demi informasi tertentu yang berhubungan dengan suatu database besar atau big data.

2.3 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 basis data, algoritma *hash based* mengumpulkan informasi mengenai $(k+1)$ -*itemset* dengan cara seluruh kemungkinan $(k+1)$ -*itemset* dihash ke dalam *hash table* dengan menggunakan fungsi *hash* (menggunakan sebuah bilangan prima untuk operasi *modulo*).

Algoritma *Hash Based* terbagi menjadi tiga bagian utama yang masing-masing bagian melakukan proses yang berbeda. Bagian pertama akan menghasilkan kandidat *1-itemset* yang disebut C_1 dan *large 1-itemset* yang disebut L_1 dari basis data. Untuk kandidat *1-itemset*, seluruh transaksi ditelusuri untuk menghitung *support count* dari *itemset* ini. Pada tahap ini *hash tree* untuk C_1 dibangun dengan tujuan mengefisienkan penghitungan *support count*. Pada bagian ini juga algoritma akan membangun *hash table* (dengan fungsi *hash*) untuk *2 itemset* yang akan berguna mengurangi banyaknya kandidat *2-itemset* C_2 .

Pada bagian kedua, kumpulan kandidat *itemset* C_k dibangkitkan berdasarkan *hash table* yang telah dibuat pada iterasi sebelumnya. Lalu ditentukan *large itemset* L_k dan mengurangi ukuran basis data untuk pembangkitan *itemset* selanjutnya. Bagian algoritma ini terbagi menjadi dua fase. Fase pertama untuk membangkitkan kandidat *k-itemset* berdasarkan *hash table*. Fase kedua akan menghitung *support* pada kandidat *itemset* dan mengurangi ukuran dari setiap transaksi.

Bagian ketiga sama seperti bagian kedua tetapi tidak menggunakan *hash table* sehingga mirip dengan algoritma apriori. Bagian kedua dilakukan selama nilai *hash* buket lebih besar dari *minimum support*. Setelah batasan ini terlewati, algoritma *hash based* diganti dengan algoritma apriori. Berikut penjelasan dari rumus *hash table* (*hash bucket*):

$$H(X,Y) = [(order\ of\ X) * penambahan\ ctr\ hash\ table + (order\ of\ y)] \bmod\ prima$$

Keterangan :

1. *Order of X* adalah perwakilan nilai X
2. Penambahan *ctr hash table* adalah nilai modulus bilangan prima, apabila terjadi *collision* nilai tersebut ditambah 1 (+ 1) sampai tidak terjadi *collision*.
3. *Order of Y* adalah perwakilan nilai Y
Prima adalah bilangan prima yang terdekat dan yang lebih besar dari jumlah kombinasi *2-itemet* (C_2).

3. HASIL DAN PEMBAHASAN

3.1 Analisa Masalah

CV. Green Uni Fruit merupakan salah satu perusahaan berkembang yang bergerak dalam bidang distributor dan importir buah – buahan dari berbagai Negara. Dalam menjalankan aktifitas operasionalnya, CV. Green Uni Fruit memiliki kendala dalam hal pemesanan buah yang dinilai masih kurang efektif dan efisien sehingga dalam proses pengerjaannya masih kurang maksimal. Dalam proses pemesanan, proses pemesanan dilakukan oleh admin dengan membuat daftar PO (*Purchasing Order*) buah. Namun yang menjadi permasalahan adalah proses pemesanan buah hanya berdasarkan asumsi dan tidak ada dasar yang digunakan pada proses pemesanan buah. Sering sekali terjadi buah yang dipesan berlebih dan tidak terjual yang menyebabkan menumpuknya stock digudang (*Over Stock*), busuk dan tidak terjual. Selain itu sering juga didapati stok buah kosong pada lapangan karena buah yang dipesan tidak memenuhi kebutuhan yang diinginkan oleh pelanggan. Kondisi ini tentu menjadi masalah yang cukup serius untuk perusahaan karena akan menimbulkan kerugian.

Berdasarkan analisa masalah diatas, maka diperlukan sebuah solusi yang diharapkan dapat mengurangi resiko kerugian tersebut. Data mining merupakan sebuah proses yang dapat dilakukan untuk mengatasi permasalahan yang dihadapi. Data mining adalah proses pengolahan data untuk mendapatkan sebuah pola atau informasi baru yang tersimpan pada data. Algoritma *hash based* merupakan teknik *hashing* yang digunakan untuk menyaring keluar *itemset* yang tidak penting dengan menentukan *frequent-itemset* dari kandidat *k-itemset*. Dengan menggunakan teknik *hashing*, kandidat *k-itemset* yang telah dipilih akan dimasukkan ke dalam *bucket* pada tabel hash. Dari *bucket* tersebut akan digunakan untuk mencari $(k+1)$ -*itemset*. Dalam skripsi ini akan dijelaskan bagaimana penerapan algoritma *hash based* yang diharapkan dapat menjadi solusi dalam permasalahan tersebut.

3.1.1 Penerapan Algoritma Hash Based

Pada data yang akan di proses dengan algoritma hash based pada data mining pemesanan buah impor pada CV. Green Uni Fruits tampak pada tabel berikut.

Tabel 1. Data Pemesanan Buah CV. Green Uni Fruits

No	Bulan 2019	Jeruk Sunkist	Longan Merah	Anggur Sunrale	Anggur Hitam Hello	Pear Madu	Apel Fuji	Apel Hijau	Kiwi Gold	Plum	Leci
1	Jan	300	600	120	340	560	600	80	320	100	100
2	Feb	700	560	500	600	580	600	380	340	100	600
3	Mar	150	250	340	100	340	120	200	600	480	160
4	Apr	350	150	560	580	320	320	180	140	300	600
5	Mei	600	200	600	300	100	340	420	140	300	100
6	Jun	560	300	100	320	300	560	100	600	580	600
7	Jul	540	580	600	580	560	120	300	340	120	600
8	Agust	150	140	300	120	100	320	400	140	200	140
9	Sept	320	600	100	540	240	300	260	580	560	480
10	Okt	100	320	120	240	560	140	150	140	600	100
11	Nov	500	340	100	260	340	400	200	340	180	560
12	Des	600	320	560	350	600	480	600	600	80	180

Dari data tersebut kemudian akan dilakukan pengolahan kembali menjadi data baru dengan format tabular yang mana jika nilai ≥ 350 akan diberi nilai 1, sedangkan nilai < 350 akan diberi nilai 0, dan untuk mempermudah penggunaan nama buah dalam perhitungan algoritma hash based maka dari nama-nama buah tersebut akan diberi inisial. Berikut ini disajikan format data tabular buah dengan inisialnya.

Tabel 2. Data Buah Dalam Bentuk Tabular

Bulan 2019	Jeruk Sunkist (A)	Longan Merah (B)	Anggur Sunrale (C)	Anggur Hitam Hello (D)	Pear Madu (E)	Apel Fuji (F)	Apel Hijau (G)	Kiwi Gold (H)	Plum (I)	Leci (J)
Jan	0	1	0	0	1	1	0	0	0	0
Feb	1	1	1	1	1	1	1	0	0	1
Mar	0	0	0	0	0	0	0	1	1	0
Apr	1	0	1	1	0	0	0	0	0	1
Mei	1	0	1	0	0	0	1	0	0	0
Jun	1	0	0	0	0	1	0	1	1	1
Jul	1	1	1	1	1	0	0	0	0	1
Agust	0	0	0	0	0	0	1	0	0	0
Sept	0	1	0	1	0	0	0	1	1	1
Okt	0	0	0	0	1	0	0	0	1	0
Nov	1	0	0	0	0	1	0	0	0	1
Des	1	0	1	1	1	1	1	1	0	0

Karena data diatas terdapat data yang missing value atau tidak ada nilai, maka data tersebut akan dibuat lagi dengan data yang bersih sehingga siap untuk dilakukan pengolahan dengan perhitungan algoritma hash based.

Tabel 3. Data Buah Telah di Cleaning

No	Bulan	Items
1	Januari	B,E,F
2	Februari	A,B,C,D,E,F,G,J
3	Maret	H,I
4	April	A,C,D,J
5	Mei	A,C,G
6	Juni	A,F,H,I,J
7	Juli	A,B,C,D,E,J
8	Agustus	G
9	September	B,D,H,I,J
10	Oktober	E,I
11	November	A,F,J
12	Desember	A,C,D,E,F,G,H

Dalam algoritma hash based langkah pertama adalah membuat L1 (large 1-itemset) dari C1 (kandidat 1-itemset) dengan menentukan nilai minsup yang diinginkan untuk mencari kandidat 2-itemset yang akan ditempatkan pada alamat tabel hash. Untuk jenis buah ini akan ditentukan nilai minsup sebesar $\geq 40\%$. Perolehan persentase nilai yang didapat pada LI (large 1-itemset).

Tabel 4. Perhitungan Nilai Support

No	Item	Frekuensi	Support
1	A	7	$7/12 * 100\% = 58,3\%$
2	B	4	$4/12 * 100\% = 33,3\%$
3	C	5	$5/12 * 100\% = 41,7\%$
4	D	5	$5/12 * 100\% = 41,7\%$
5	E	5	$5/12 * 100\% = 41,7\%$
6	F	5	$5/12 * 100\% = 41,7\%$
7	G	4	$4/12 * 100\% = 33,3\%$
8	H	4	$4/12 * 100\% = 33,3\%$
9	I	4	$4/12 * 100\% = 33,3\%$
10	J	6	$6/12 * 100\% = 50\%$

Dari tabel diatas didapatkan nilai support, item yang tidak memenuhi nilai minimum support sebesar 40% maka item tersebut tidak digunakan untuk proses. Karena algoritma hash based menggunakan tabel hash yang memiliki nilai alamat tertentu, maka dari nama buah tersebut digunakan perwakilan nilai alamat tertentu, maka dari nama buah tersebut digunakan perwakilan nilai yang nantinya akan diproses sehingga menghasilkan nilai kunci yang akan menempatkan kombinasi buah tersebut pada nilai tabel hash yang tersedia.

Tabel 5. Hash Based

No	Item	Alamat
1	A	1
2	C	2
3	D	3
4	E	4
5	F	5
6	J	6

Pada proses sebelumnya telah ditentukan L1 (large 1-itemset), dari item yang terdapat pada L1 (large 1-itemset) tersebut kemudian akan dikombinasikan menjadi C2 (kandidat 2-itemset) sejumlah 15 kombinasi yaitu : (AC), (AD), (AE), (AF), (AJ), (CD), (CE), (CF), (CJ), (DE), (DF), (DJ), (EF), (EJ), (FJ) Untuk memperoleh nilai alamat hash maka digunakan rumus hash bucket :

$$H(X,Y) = [(order\ of\ X) * penambahan\ ctr\ hash\ table + (order\ of\ y)] \text{ mod prima}$$

Perhitungan ini dimulai dengan perkalian 17 dengan mode 17 yang mana angka tersebut diperoleh dari bilangan prima yang terdekat dan yang lebih besar dari 15 (jumlah C2 (kombinasi 2-itemset)). Berikut perhitungan dengan menggunakan mod 17:

$$\begin{aligned} h(AC) &= [(1)*17+(2)] \text{ mod } 17 = 2 \\ h(AD) &= [(1)*17+(3)] \text{ mod } 17 = 3 \\ h(AE) &= [(1)*17+(4)] \text{ mod } 17 = 4 \\ h(AF) &= [(1)*17+(5)] \text{ mod } 17 = 5 \\ h(AJ) &= [(1)*17+(6)] \text{ mod } 17 = 6 \\ h(CD) &= [(2)*17+(3)] \text{ mod } 17 = 3 \\ h(CE) &= [(2)*17+(4)] \text{ mod } 17 = 4 \\ h(CF) &= [(2)*17+(5)] \text{ mod } 17 = 5 \\ h(CJ) &= [(2)*17+(6)] \text{ mod } 17 = 6 \\ h(DE) &= [(3)*17+(4)] \text{ mod } 17 = 4 \\ h(DF) &= [(3)*17+(5)] \text{ mod } 17 = 5 \\ h(DJ) &= [(3)*17+(6)] \text{ mod } 17 = 6 \\ h(EF) &= [(4)*17+(5)] \text{ mod } 17 = 5 \\ h(EJ) &= [(4)*17+(6)] \text{ mod } 17 = 6 \\ h(FJ) &= [(5)*17+(6)] \text{ mod } 17 = 6 \end{aligned}$$

Pada hasil perhitungan hash diatas masih terdapat alamat nilai hash yang masih sama atau collision. Maka harus dilakukan perhitungan dengan menggunakan nilai mod yang baru. Karena dengan menggunakan modulus 17 sampai perhitungan ke 15 kali masih terdapat nilai yang terjadi *collision* antar item C2 (kandidat 2-itemset), maka modulus yang digunakan diganti dengan bilangan modulus 19, karena urutan modulus setelah 17 adalah 19. Dan nilai ctr *hash table* ditambah +1

$$\begin{aligned} h(AC) &= [(1)*18+(2)] \text{ mod } 19 = 1 \\ h(AD) &= [(1)*18+(3)] \text{ mod } 19 = 2 \\ h(AE) &= [(1)*18+(4)] \text{ mod } 19 = 3 \\ h(AF) &= [(1)*18+(5)] \text{ mod } 19 = 4 \end{aligned}$$

$$\begin{aligned} h(AJ) &= [(1)*18+(6)] \text{ mod } 19 = 5 \\ h(CD) &= [(2)*18+(3)] \text{ mod } 19 = 1 \\ h(CE) &= [(2)*18+(4)] \text{ mod } 19 = 2 \\ h(CF) &= [(2)*18+(5)] \text{ mod } 19 = 3 \\ h(CJ) &= [(2)*18+(6)] \text{ mod } 19 = 4 \\ h(DE) &= [(3)*18+(4)] \text{ mod } 19 = 1 \\ h(DF) &= [(3)*18+(5)] \text{ mod } 19 = 2 \\ h(DJ) &= [(3)*18+(6)] \text{ mod } 19 = 3 \\ h(EF) &= [(4)*18+(5)] \text{ mod } 19 = 1 \\ h(EJ) &= [(4)*18+(6)] \text{ mod } 19 = 2 \\ h(FJ) &= [(5)*18+(6)] \text{ mod } 19 = 1 \end{aligned}$$

Pada hasil perhitungan hash diatas masih terdapat alamat nilai hash yang masih sama atau *collision*. Maka harus dilakukan perhitungan dengan menggunakan nilai mod yang baru. Karena dengan menggunakan modulus 19 sampai perhitungan ke 15 kali masih terdapat nilai yang terjadi *collision* antar item C2 (kandidat 2-itemset), maka modulus yang digunakan diganti dengan bilangan modulus 23, karena urutan modulus setelah 19 adalah 23. Dan nilai ctr *hash table* ditambah +1

$$\begin{aligned} h(AC) &= [(1)*19+(2)] \text{ mod } 23 = 21 \\ h(AD) &= [(1)*19+(3)] \text{ mod } 23 = 22 \\ h(AE) &= [(1)*19+(4)] \text{ mod } 23 = 0 \\ h(AF) &= [(1)*19+(5)] \text{ mod } 23 = 1 \\ h(AJ) &= [(1)*19+(6)] \text{ mod } 23 = 2 \\ h(CD) &= [(2)*19+(3)] \text{ mod } 23 = 18 \\ h(CE) &= [(2)*19+(4)] \text{ mod } 23 = 19 \\ h(CF) &= [(2)*19+(5)] \text{ mod } 23 = 20 \\ h(CJ) &= [(2)*19+(6)] \text{ mod } 23 = 21 \\ h(DE) &= [(3)*19+(4)] \text{ mod } 23 = 15 \\ h(DF) &= [(3)*19+(5)] \text{ mod } 23 = 16 \\ h(DJ) &= [(3)*19+(6)] \text{ mod } 23 = 17 \\ h(EF) &= [(4)*19+(5)] \text{ mod } 23 = 12 \\ h(EJ) &= [(4)*19+(6)] \text{ mod } 23 = 13 \\ h(FJ) &= [(5)*19+(6)] \text{ mod } 23 = 9 \end{aligned}$$

Pada hasil perhitungan hash diatas masih terdapat alamat nilai hash yang masih sama atau *collision*. Maka harus dilakukan perhitungan dengan menggunakan nilai mod yang baru. Karena dengan menggunakan modulus 23 sampai perhitungan ke 15 kali masih terdapat nilai yang terjadi *collision* antar item C2 (kandidat 2-itemset), maka modulus yang digunakan diganti dengan bilangan modulus 29, karena urutan modulus setelah 23 adalah 29. Dan nilai ctr *hash table* ditambah +1

$$\begin{aligned} h(AC) &= [(1)*20+(2)] \text{ mod } 29 = 22 \\ h(AD) &= [(1)*20+(3)] \text{ mod } 29 = 23 \\ h(AE) &= [(1)*20+(4)] \text{ mod } 29 = 24 \\ h(AF) &= [(1)*20+(5)] \text{ mod } 29 = 25 \\ h(AJ) &= [(1)*20+(6)] \text{ mod } 29 = 26 \\ h(CD) &= [(2)*20+(3)] \text{ mod } 29 = 14 \\ h(CE) &= [(2)*20+(4)] \text{ mod } 29 = 15 \\ h(CF) &= [(2)*20+(5)] \text{ mod } 29 = 16 \\ h(CJ) &= [(2)*20+(6)] \text{ mod } 29 = 17 \\ h(DE) &= [(3)*20+(4)] \text{ mod } 29 = 6 \\ h(DF) &= [(3)*20+(5)] \text{ mod } 29 = 7 \\ h(DJ) &= [(3)*20+(6)] \text{ mod } 29 = 8 \\ h(EF) &= [(4)*20+(5)] \text{ mod } 29 = 27 \\ h(EJ) &= [(4)*20+(6)] \text{ mod } 29 = 28 \\ h(FJ) &= [(5)*20+(6)] \text{ mod } 29 = 19 \end{aligned}$$

Dari perhitungan nilai hash diatas sudah tidak ada lagi item yang memiliki nilai alamat yang sama atau *collision*. Maka proses selanjutnya adalah menyusun alamat hash dan *basket count*

Tabel 6. Basket Count

No	Alamat	Item	Basket Count
1	6	[DE]	3
2	7	[DF]	2
3	8	[DJ]	3
4	14	[CD]	4

No	Alamat	Item	Basket Count
5	15	[CE]	3
6	16	[CF]	2
7	17	[CJ]	3
8	19	[FJ]	3
9	22	[AC]	5
10	23	[AD]	4
11	24	[AE]	3
12	25	[AF]	3
13	26	[AJ]	5
14	27	[EF]	3
15	28	[EJ]	2

Kemudian setelah didapatkan perhitungan *basket count*, selanjutnya untuk melakukan perhitungan nilai support dalam menentukan proses selanjutnya

Tabel 7. Kombinasi 3 Itemset

No	Alamat	Item	Basket Count	Support
1	6	[DE]	3	25%
2	7	[DF]	2	16,7%
3	8	[DJ]	3	25%
4	14	[CD]	4	33%
5	15	[CE]	3	25%
6	16	[CF]	2	16,7%
7	17	[CJ]	3	25%
8	19	[FJ]	3	25%
9	22	[AC]	5	41,7%
10	23	[AD]	4	33,3%
11	24	[AE]	3	25%
12	25	[AF]	3	25%
13	26	[AJ]	5	41,7%
14	27	[EF]	3	25%
15	28	[EJ]	2	16,7%

Dengan nilai minimum support sebesar 40% yang telah ditentukan maka didapatkan bahwa hanya terdapat 2 kombinasi item yang memenuhi nilai minimum support yaitu

Tabel 8. Kombinasi 2 Item Set

No	Alamat	Item	Basket Count	Support
1	22	[AC]	5	41,7%
2	26	[AJ]	5	41,7%

Kemudian dari item yang memenuhi nilai minimum support kemudian dibentuk kombinasi item set dengan 3 item set baru

Tabel 9. Support Hash Based

No	Kombinasi Item	Basket Count	Support
1	[ACJ]	3	25%

Setelah menghitung nilai support kemudian menghitung nilai confidence dari kombinasi item AC→J.

Tabel 10. Confidence Hash Based

No	Kombinasi Item	Basket Count	Support	Confidence
1	[AC→J]	3	25%	$3/5 * 100\% = 60\%$

Maka dari proses yang telah dilakukan pada algoritma Hash Based didapatkan bahwa 3 item yang menjadi hasil. Item tersebut yaitu ; A=Jeruk Sunkist, C=Anggur Sunrale dan J=Leci yang menjadi prioritas buah pemesanan.

4. KESIMPULAN

Berdasarkan pembahasan yang telah dilakukan penulis pada penelitian ini maka ada beberapa hal yang dijadikan sebagai kesimpulan pada penelitian ini Pada penelitian yang dilakukan pengolahan data pemesanan di CV. Green Uni Fruit memanfaatkan proses pada data mining untuk mendapatkan informasi baru terkait dengan pola pemesanan buah. Dengan menggunakan algoritma hash based didapatkan hasil penelitian berupa sebuah pola *asosiasi* buah Jeruk Sunkist,

Anggur Sunrale dan Leci merupakan kombinasi buah yang diprioritaskan untuk pemesanan dengan nilai *support* sebesar 25% dan *confidence* 60%

REFERENCES

- [1] F. Ramadhan, "Implementasi Algoritma Hash Based Terhadap Aturan Asosiasi untuk Menentukan Frequent Itemset Study Kasus Rumah Makan Seafood 'Kita,'" *Semin. Nas. Teknol. Inf. dan Multimed.*, pp. 97–102, 2017.
- [2] 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, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTISI>.
- [3] T. D. Prakoso, I. Ernawati, and H. B. Seta, "Penemuan Pola Asosiasi Pada Data Restoran Menggunakan Algoritma Hash Based," *Semin. Nas. Mhs. Ilmu Komput. dan Apl.*, pp. 71–80, 2020.
- [4] O. S. A. Destiyati and E. Aribowo, "Analisis Perbandingan Algoritma Apriori Dan Algoritma Hash Based Pada Market Basket Analysis Di Apotek UAD," *JSTIE (Jurnal Sarj. Tek. Inform.)*, vol. 3, no. 1, pp. 1–10, 2015, doi: 10.12928/jstie.v3i1.2896.
- [5] R. D. Yulanda, S. Wahyuningsih, and F. D. T. Amijaya, "Association rules with apriori algorithm and hash-based algorithm," *J. Phys. Conf. Ser.*, vol. 1277, no. 1, pp. 1–9, 2019, doi: 10.1088/1742-6596/1277/1/012048.
- [6] E. Prasetyo, *Data Mining: Konsep Dan Aplikasi Menggunakan Matlab*. Yogyakarta: CV. Andi Offset, 2012.
- [7] D. Nofriansyah and G. W. Nurcahyo, *Algoritma Data Mining Dan Pengujiannya*. Yogyakarta: Deepublish, 2017.
- [8] N. S. Pinem and D. P. Utomo, "Implementasi Fuzzy Logic Dengan Infrensi Tsukamoto Untuk Prediksi Jumlah Kemasan Produksi (Studi Kasus: PT. Sinar Sosro Medan)," *Pelita Inform. Inf. dan Inform.*, vol. 9, no. 1, pp. 56–60, 2020.
- [9] D. P. Utomo and Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. MEDIA Inform. BUDIDARMA*, vol. 4, no. 2, pp. 437–444, 2020.
- [10] D. P. Utomo, P. Sirait, and R. Yunis, "Reduksi Atribut Pada Dataset Penyakit Jantung dan Klasifikasi Menggunakan Algoritma C5. 0," *J. MEDIA Inform. BUDIDARMA*, vol. 4, no. 4, pp. 994–1006, 2020.
- [11] B. S. Pranata and D. P. Utomo, "Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor (Study Kasus Bengkel Sinar Service)," *Bull. Inf. Technol.*, vol. 1, no. 2, pp. 83–91, 2020.
- [12] D. P. Utomo and B. Purba, "Penerapan Datamining pada Data Gempa Bumi Terhadap Potensi Tsunami di Indonesia," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. 1, pp. 846–853, 2019.
- [13] R. Amelia and D. P. Utomo, "ANALISA POLA PEMESANAN PRODUK MODERN TRADE INDEPENDENT DENGAN MENEREPAKAN ALGORITMA FP. GROWTH (STUDI KASUS: PT. ADAM DANI LESTARI)," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 416–423, 2019.