

GUI R UNTUK ANALISIS KERANJANG BELANJA DENGAN ALGORITMA APRIORI PADA SUATU PERUSAHAAN E-COMMERCE

Ryan Anugrah¹, Tatik Widiharih², Sugito³

^{1,2,3}Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro

e-mail: ryan18ln@gmail.com

ABSTRACT

Technological developments help people live easier. One of the technological developments is being able to trade digitally or it can be called e-commerce. To increase revenue, e-commerce companies collect consumer sales history data that can be analyzed and obtain information about consumer habits. One of the analyzes that can be used is shopping basket analysis which aims to find a pattern in transaction data. In data processing and analysis is done using the R program computation and GUI R is made with a recommendation system simulation. The results of the shopping cart analysis produce as many as 22 rules using a minimum support of 0.06 and a confidence of 0.5. The greater the support value, the more often the product or rule is purchased by consumers from all data transactions and vice versa. Meanwhile, the greater the trust value, the more often the products purchased under the regulation are purchased together. Thus, the information can be used to help carry out promotions to increase sales by the company.

Keywords : Data Mining, Association Rule, Market Basket Analysis, Algorithm Apriori, E-commerce

1. PENDAHULUAN

Ada berbagai macam teknologi yang telah berkembang hingga saat ini. Salah satunya adalah internet, internet merupakan teknologi yang membuat hilangnya jarak yang ada dari manusia satu yang ada di belahan bumi dengan yang ada di belahan bumi lainnya. Selain hilangnya jarak komunikasi antar manusia, internet juga membantu manusia dalam perdagangan sehingga memiliki pasar yang luas. Perdagangan tersebut biasa disebut sebagai *electronic commerce*. Perdagangan elektornik atau *E-commerce* adalah hasil teknologi informasi yang saat ini sedang berkembang dengan begitu cepat terhadap pertukaran barang, jasa dan informasi melalui sistem elektornik seperti: internet, televisi, dan jaringan komputer lainnya (Limbong, 2019).

Suatu perusahaan pasti melakukan analisis untuk meningkatkan pemasukkan dan melakukan promosi untuk membuat lebih banyak pelanggan yang membeli atau berbelanja pada perusahaan tersebut. Begitu juga dengan yang dilakukan oleh perusahaan e-commerce, dari berbagai transaksi yang telah terjadi tersimpan data penjualan produk. Berdasarkan data penjualan produk tersebut dapat dilakukan berbagai analisis yang dapat membantu mengembangkan usahanya. Salah satunya ialah analisis keranjang belanja.

Analisis keranjang belanja atau Market basket analysis adalah suatu teknik matematis yang dipakai secara umum oleh profesional pemasaran untuk mengungkap persamaan atau hubungan keterikatan antar produk maupun antar kumpulan produk. Market basket analysis menentukan produk apa yang dibeli bersamaan. Setelah perusahaan mengetahui ketika satu produk akan dibeli bersamaan dengan produk lainnya, maka perusahaan dapat memasarkan dan menjadikan pembeli produk tertentu sebagai target prospek untuk produk lainnya. Hal ini akan meningkatkan tingkat efektivitas pemasaran dan taktik penjualan dengan menggunakan data konsumen yang tersedia (Haryanto, 2011).

Data-data penjualan dianalisis untuk mencari pola keterikatan untuk tiap-tiap barang. Setelah pola keterikatan tiap barang didapat, kemudian dicari tingkat keterikatannya. Tingkat keterikatan dari tiap barang dapat berbeda-beda tergantung dari data yang ada.

Setiap keterikatan dari 1 barang dengan barang lainnya dikenal dengan istilah rule atau aturan. Jika analisis ini dilakukan secara manual, tentu membutuhkan waktu dan tenaga yang cukup banyak. Sehingga analisisnya dilakukan menggunakan komputasi salah satu program yang dapat digunakan untuk melakukan analisis ini yaitu program R.

Dalam analisis akan dibetuk suatu GUI pada program R yang bertujuan untuk memudahkan bagi pengguna yang ingin menganalisis data transaksinya menggunakan analisis keranjang belanja. Selain itu pada GUI yang akan dibuat terdapat simulasi rekomendasi produk. Simulasi tersebut untuk melihat satu produk dapat direkomendasikan dengan produk-produk lain yang memiliki kemungkinan untuk dibelanjakan secara bersamaan yang diperoleh dari hasil analisis keranjang belanja.

Berdasarkan uraian diatas maka dari informasi yang ada akan dilakukan analisis berupa analisis keranjang belanja. Analisis dilakukan untuk menggali informasi - informasi tentang kebiasaan konsumen dalam membeli produk yang ditransaksikan dari data yang ada. Sehingga dari analisis yang dilakukan nantinya diharapkan dapat membantu perusahaan dengan informasi yang didapatkan.

2. TINJAUAN PUSTAKA

Menurut Ruth (2003), definisi e-commerce adalah penggunaan komunikasi elektronik dan teknologi informasi dalam transaksi bisnis untuk menciptakan, mentransformasikan, dan mendefinisikan kembali hubungan penciptaan harga di antara organisasi, dan antara organisasi dan individu.

Mengacu pada penelitian Mougayar (dalam Soliman dan Youssef, 2003) mengenai ukuran dalam menentukan nilai suatu bisnis, E-Commerce mempunyai beberapa kelebihan nilai bisnis yaitu:

1. Reducing cost
2. Process simplification
3. Improving customer service
4. Generating new revenue
5. Taking faster decisions

Perilaku konsumen adalah aktivitas konsumen dalam memutuskan untuk membeli, menggunakan, serta mengkonsumsi membeli barang dan jasa termasuk dalam faktor-faktor pelanggan yang dapat menimbulkan keputusan mereka apakah akan membeli dan menggunakan produk. Setiap pelanggan memiliki kebutuhan dan kecenderungan yang berbeda pula karena memiliki perilaku yang berbeda dalam memenuhi hal-hal tersebut. Namun, jika terjadi perilaku yang berbeda memenuhi kebutuhan mereka, mereka masih memiliki beberapa kesamaan, salah satunya berkeinginan untuk memaksimalkan kepuasan mereka dalam mengkonsumsi produk atau layanan yang diperlukan (Kurniawan et al., 2018).

Berdasarkan pada sudut pandang fungsi statistik dalam sebuah penelitian, statistik deskriptif (deduktif) dapat diartikan sebagai statistik yang berfungsi untuk mendeskripsikan fenomena-fenomena terteliti berdasarkan data yang terkumpul. Statistik deskriptif tidak bermaksud menarik sebuah kesimpulan, namun hanya terbatas pada penyajian data yang telah terkumpul dan diolah/disusun dalam bentuk tabel, grafik (diagram), tendensi sentral dan variasi (variabilitas), agar dapat memberi gambaran yang teratur, ringkas, dan jelas mengenai data suatu peristiwa atau keadaan (Mundir, 2012).

Penyusunan dan pengklasifikasian data yang telah terkumpul, sebagaimana yang kita ketahui bersama, merupakan cara yang perlu ditempuh agar mudah diamati, diinterpretasikan dan diproses lebih lanjut. Atas dasar kaidah ilmu statistik, upaya itu

dilakukan dengan menyusun suatu distribusi frekuensi (frequency distribution). Distribusi frekuensi (atau sering disebut dengan tabel distribusi frekuensi) merupakan penyusunan data menurut besar kecilnya data itu dalam beberapa kelas tertentu. Melalui distribusi frekuensi, data akan dikelompokkan ke dalam beberapa golongan dengan menerapkan kaidah tertentu.

Menurut Santoso dan Hamdani (2007), Penyajian data dalam bentuk diagram ini bertujuan untuk memudahkan para pembaca terutama orang awam yang tidak biasa mengambil informasi dari data mentah yang belum diolah. Penyajian data dalam bentuk diagram dapat ditentukan sesuai dengan kebutuhan informasi yang diperlukan.

Menurut Agrawal, R (dalam Tampubolon et al., 2013), data mining diartikan sebagai menambang data atau upaya untuk menggali informasi yang berharga dan berguna pada database yang sangat besar. Menurut Linoff dan Berry (2011), data mining adalah proses bisnis untuk mengeksplorasi sejumlah data yang besar untuk menemukan pola dan aturan yang bermakna. Menurut Hermawati (2013), data mining adalah proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer (mahine learning) unuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis.

Tujuan dari data mining yaitu, untuk mendapatkan pengetahuan yang masih tersembunyi di dalam bongkahan data (Susanto dan Suryadi, 2010). Di dalam data mining terdapat istilah pattern recognition atau pengenalan pola yang tepat digunakan karena pengetahuan yang hendak digali memang berbentuk pola-pola yang mungkin juga masih perlu digali dari dalam bongkahan data (Susanto dan Suryadi, 2010).

Aturan asosiasi digunakan untuk mendeteksi hubungan umum antara item, sehingga membuat analisis keranjang pasar sebagai aplikasi praktis inferensi dan penggunaan aturan (Berry dan Browne, 2006). Analisis keranjang belanja tidak mengacu pada suatu teknik melainkan sesuai dengan namanya yaitu, merujuk pada serangkaian masalah bisnis yang terkait dengan pemahaman data transaksi penjualan. Aplikasi dari analisis ini telah diperluas ke banyak domain berbeda. Misalnya, aplikasi analisis keranjang pasar yang lebih baru adalah untuk memahami bagian-bagian situs web yang dikunjungi pelanggan (Linoff dan Berry, 2011).

Aturan asosiasi pertama kali diperkenalkan oleh Agrawal et al. pada tahun 1993 sebagai sarana untuk menentukan hubungan di antara set item dalam database. Aturan asosiasi, seperti clustering, adalah bentuk unsupervised learning dan telah diterapkan ke banyak bidang seperti bisnis ritel, web mining, dan text mining (Berry dan Browne, 2006). Aturan asosiasi yang berbentuk “if ... then ...” atau “jika ... maka ...” merupakan bentuk umum yang akan dihasilkan dari aturan asosiasi (Susanto dan Suryadi, 2010).

Misalkan $I = \{i_1, i_2, \dots, i_d\}$ menjadi himpunan semua item dalam data keranjang belanja dan $T = \{t_1, t_2, \dots, t_N\}$ menjadi himpunan semua transaksi. Setiap transaksi t_i berisi subset item yang dipilih dari I . Dalam analisis asosiasi, kumpulan nol atau lebih item disebut itemset. Jika itemset berisi k item, itu disebut k -itemset. Misalnya, {Tomat, Bawang putih, Bawang merah} adalah contoh 3-itemset. Set null (atau kosong) adalah itemset yang tidak mengandung item apa pun (Tan et al., 2004).

Menurut Kusriani (dalam Listriani et al., 2018), metodologi dasar aturan asosiasi terbagi menjadi dua tahap, yaitu :

1. Analisis pola frekuensi tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Nilai support (penunjang) yaitu persentase item item atau kombinasi item yang ada pada database. Nilai support sebuah item diperoleh dengan rumus berikut:

$$\text{Support (A)} = \frac{\text{Jumlah transaksi mengandung A}}{\text{Jumlah transaksi}}$$

Sedangkan nilai dari support 2 item diperoleh dari rumus berikut :

$$\text{Support (A} \rightarrow \text{B)} = \frac{\text{Jumlah transaksi mengandung A dan B}}{\text{Jumlah transaksi}}$$

2. Pembentukan aturan asosiatif

Setelah semua pola frekuensi tinggi ditemukan, maka dicari aturan asosiatif yang memenuhi syarat minimum untuk confidence dengan menghitung confidence aturan asosiatif “ Jika A maka B” = (A → B).

$$\begin{aligned} \text{Confidence (A} \rightarrow \text{B)} &= P(\text{B}|\text{A}) \\ &= \frac{\text{Jumlah transaksi yang ada A dan B}}{\text{Jumlah transaksi yang ada A}} \end{aligned}$$

Menurut Linoff dan Berry, 2011, teknik aturan asosiasi harus memisahkan yang kuat dari yang lemah, jadi dari sejumlah besar aturan yang mungkin, hanya yang terbaik yang dipilih. Tiga metode tradisional untuk mengukur kebaikan aturan asosiasi adalah support, confidence, dan lift.

- Support mengukur jumlah atau proporsi transaksi yang mengandung semua item dalam aturan.
- Confidence mengukur seberapa baik aturan dalam memprediksi sisi kanan (consequent), dengan membandingkan seberapa sering sisi kanan (consequent) muncul ketika kondisi di sisi kiri (antecedent).
- lift adalah ukuran untuk melihat seberapa baik performa aturan tersebut. Lift dihitung sebagai rasio confidence aturan dengan prevalensi sisi kanan. Atau dapat di tulis dengan rumus berikut :

$$\text{Lift (A} \rightarrow \text{B)} = \frac{\text{confidence (A} \rightarrow \text{B)}}{\text{Jumlah transaksi B}}$$

Jika lift lebih besar dari 1, aturan yang dihasilkan lebih baik dalam memprediksi hasil daripada menebak apakah sisi kanan aturan ada berdasarkan frekuensi item dalam data. Jika lift kurang dari 1, aturan berjalan lebih buruk daripada menebak yang diinformasikan

Sistem rekomendasi membantu pengguna untuk mengidentifikasi produk yang sesuai dengan kebutuhan, kesenangan, dan keinginan user. Sistem rekomendasi akan membimbing user untuk menemukan produk yang relevan dan berguna dari banyaknya produk yang tersedia (Prasetya, 2017).

Menurut Tang et al. dalam Prasetya, (2017) sejak pertama kali ditemukan pada tahun 1990, ada banyak penelitian mengenai sistem rekomendasi dilakukan. Sistem rekomendasi mulai diaplikasikan ke berbagai bidang dengan metode yang berbeda seperti Content Based, Collaborative Filtering, dan Hybrid. Sistem rekomendasi yang menggunakan metode Content Based menggunakan kesamaan produk untuk ditawarkan kepada pengguna. Namun, metode Content Based memiliki kelemahan yaitu ketika fitur konten yang tersedia terbatas, maka akurasi rekomendasi yang dihasilkan cukup rendah (Yuan et al. dalam Prasetya, 2017).

Collaborative Filtering adalah metode yang paling sering digunakan untuk membangun system rekomendasi. Metode ini bergantung pada Riwayat pemilihan atau riwayat penilaian (Su dan Khoshgoftaar dalam Prasetya, 2017). Metode Hybrid menggabungkan metode Content Based dan Collaborative Filtering untuk menghasilkan rekomendasi yang lebih baik (Tang et al. dalam Prasetya, 2017).

Hubungan antara aturan asosiasi dan filter kolaboratif alami karena masalah aturan asosiasi pertama kali diusulkan dalam konteks menemukan hubungan antara data supermarket. Aturan asosiasi secara alami didefinisikan atas data biner, meskipun pendekatan dapat diperluas ke data kategorikal dan numerik dengan mengubah tipe data ini menjadi data biner (Agrawal, 2016).

R Analytics (atau bahasa pemrograman R) adalah perangkat lunak sumber terbuka gratis (open-source) yang digunakan untuk komputasi statistik. Bahasa ini dibangun khusus untuk,

dan digunakan secara luas oleh olah data statistik serta penggalian data (Prasetyo et al., 2020). Pemanfaatan R dikalangan para peneliti dan pengajar statistika pada umumnya, kalah populer dengan software software berbayar seperti SPSS, MINITAB. Salah satu penyebabnya adalah karena program R (sebagaimana kebanyakan open source) sebagian besar menggunakan pendekatan skrip atau CLI (Command Line Interface), yang lebih cocok untuk para pengembang, bukan para pengguna statistika. Sebenarnya usaha untuk membuat kemampuan R bisa diakses melalui menu grafis, GUI (Graphical User Interface) telah dimulai oleh beberapa pengembang untuk berbagai flatform (Tirta, 2015).

Menurut Tirta (2015), salah satu package R yang dapat digunakan untuk membangun GUI adalah Shiny. Shiny merupakan interface yang memungkinkan pengguna membuat laman web (web pages) interaktif sehingga kemampuan R yang pada dasarnya bersifat CLI bisa diakses melalui menu web secara GUI web.

3. METODE PENELITIAN

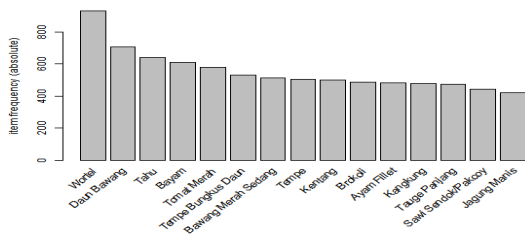
Data yang akan digunakan pada penelitian ini adalah data transaksi jual beli. Jenis data ini adalah data sekunder yang diambil dari perusahaan e-commerce. Transaksi yang diambil mulai dari bulan Mei hingga Oktober 2020. Pada penelitian ini terdapat 2 variabel utama yang digunakan untuk melakukan analisis keranjang belanja yaitu, data ID transaksi dan data produk yang dipesan oleh konsumen. Pengolahan data menggunakan software R. Berikut adalah Langkah-langkah pengolahan data dalam penelitian ini :

1. Memilih 2 variabel yang sudah ditentukan, dari database transaksi.
2. Membentuk aturan-aturan asosiasi dengan menggunakan algoritma apriori
3. Membuat plot yang berisi berbagai nilai support dan confidence sehingga akan lebih mudah menentukan nilai minimum support dan nilai minimum confidence
4. Menentukan dan memilih nilai minimum support dan minimum confidence yang akan digunakan
5. Membentuk aturan asosiasi yang memenuhi nilai minimum support dan minimum confidence
6. Melakukan interpretasi hasil dari aturan-aturan asosiasi yang telah terbentuk dari metode analisis keranjang belanja

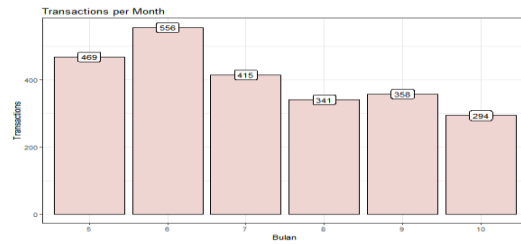
4. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini merupakan data transaksi pembelian produk pada aplikasi e-commerce. Terdapat berbagai variabel pada data tersebut, yang utama digunakan dalam analisis keranjang belanja hanya variabel ID transaksi dan produk yang dipesan oleh konsumen. Namun, untuk menambah informasi ada variabel lain yang akan digunakan yaitu variabel tanggal dan waktu. Dari variabel tersebut akan dilakukan deskriptif secara visual. Berikut merupakan statistik deskriptif dari data yang digunakan. Ada 2433 transaksi dan 380 produk yang terjual dengan periode data bulanan terhitung dari bulan Mei sampai dengan Oktober 2020.

Gambar 1 merupakan grafik 15 produk dari 380 produk yang paling banyak dibeli oleh konsumen. Wortel merupakan produk dengan penjualan paling laris dilanjutkan daun bawang dan tahu. Dari 15 top produk hampir semuanya merupakan produk sayuran kecuali ayam fillet, sehingga dapat dinyatakan sayuran merupakan pembelian paling favorit dari konsumen.



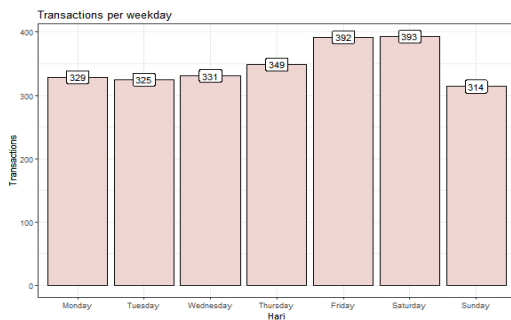
Gambar 1. Grafik pembelian produk



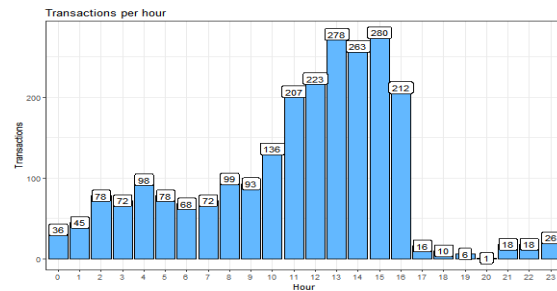
Gambar 2. Grafik transaksi tiap bulan

Gambar 2 merupakan grafik dari total transaksi tiap bulannya. Transaksi tertinggi terjadi pada bulan ke-6 yaitu bulan Juni dan terendah pada bulan ke-10 yaitu bulan Oktober. Dari grafik tersebut terdapat penurunan transaksi dari bulan Juni ke bulan Oktober, penurunan hampir 50% dari transaksi tertinggi.

Gambar 3 merupakan grafik total transaksi selama 6 bulan yang dijumlahkan berdasarkan hari dari Senin hingga Minggu. Transaksi yang pada tiapharinya memiliki jumlah yang tidak terlalu berbeda, tetapi pada hari jumat dan sabtu adalah hari yang tertinggi dari hari lainnya.



Gambar 3. Grafik transaksi tiap hari



Gambar 4. Grafik transaksi tiap jam

Gambar 4 merupakan grafik dari total transaksi selama 6 bulan yang dijumlahkan dan dibagi berdasarkan jam transaksi tersebut dilakukan. Transaksi terbanyak terjadi pada siang hari pada pukul 11 hingga pukul 16 sedangkan pada malam hari mulai pukul 17 hingga 23 sangat sedikit transaksi yang terjadi.

Inti dari analisis keranjang belanja adalah membangkitkan rules atau aturan-aturan yang berisikan ketertarikan antara satu produk dengan produk lainnya yang didapat dari history transaksi para konsumen. Data transaksi pada lampiran 1 yang digunakan hanya variabel order dan item bought atau produk. Aturan – aturan dari data tersebut akan dibangkitkan menggunakan algoritma apriori. Dalam pencarian rules menggunakan algoritma apriori sebelumnya diperlukan pemilihan minimum support dan confidence yang akan dipilih melalui percobaan dari suatu grafik yang menghasilkan plot yang berisi berapa rules yang didapat jika menggunakan support dan confidence tertentu. Support dan Confidence akan dipilih sesuai keinginan peneliti yang hasilnya akan dijadikan minimum support dan confidence untuk pencarian rules.

Frequent itemset adalah semua itemset yang memenuhi ambang batas support atau minimum support. Itemset merupakan sebuah himpunan yang berisi produk baik itu 1 produk ataupun gabungan 2 atau lebih produk. Misalnya, {Wortel} adalah contoh itemset yang berisi 1 produk sedangkan {Wortel, Tahu} adalah contoh itemset 2 produk yang merupakan gabungan dari 1 itemset {Wortel} dan 1 itemset {Tahu}. Suatu produk dapat

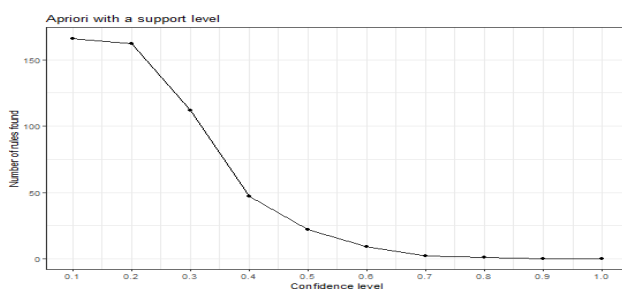
dijadikan itemset bila produk tersebut terdapat pada salah satu transaksi dan 2 produk yang dapat digabungkan pada itemset harus berada bersamaan pada salah satu transaksi.

Pada setiap itemset mempunyai nilai support yang berbeda tergantung berapa banyak transaksi yang mengandung suatu produk. Nilai support 1-itemset dan 2-itemset dengan terdapat produk yang sama akan memiliki nilai support yang lebih besar pada 1-itemset dari pada 2-itemset, hal ini merupakan salah satu sifat dari algoritma apriori. Algoritma apriori akan mencari seluruh itemset yang ada pada data transaksi sehingga tidak ada lagi itemset yang terbentuk dari data. Kombinasi dalam itemset tersebut kemudian akan dipilih yang termasuk frequent itemset dengan cara memilih itemset yang memiliki nilai support yang sama atau diatas ambang batas nilai support. Dikarekan besarnya data sehingga proses tersebut akan dibantu menggunakan program R.

Aturan asosiasi terbentuk dari itemset yang berisi 2 produk atau lebih. Misalnya, itemset {Wortel, Tahu} memiliki nilai support 0.12988 atau 12.99 %, nilai tersebut berarti terdapat 12.99% produk tersebut dibeli bersamaan dari seluruh transaksi. Dari itemset tersebut dapat dijadikan aturan asosiasi sebagai {Wortel} → {Tahu} yang dapat diartikan jika beli wortel maka beli tahu dengan nilai support 12.99 %. Permisalan tersebut diambil dari gui pada menu analisis keranjang belanja dengan hasil tabel pada lampiran 2.

Dalam memilih minimum support dan confidence dalam pencarian rules dilain sesuai keinginan peneliti tetapi juga diharapkan tidak terlalu besar dan terlalu kecil untuk memilih support dan confidencenya. Karena jika terlalu besar ditakutkan rules yang didapat terlalu sedikit dan jika terlalu kecil maka rules yang didapat terlalu banyak dan rules nya memiliki ukuran yang kecil. Sehingga rules yang ingin dicari ialah rules yang memiliki support dan confidence yang tidak begitu kecil dan tidak terlalu sedikit rules yang didapatkan.

Untuk memudahkan memilih support dan confidence yang sesuai maka digunakan suatu plot yang menampilkan banyak rules pada setiap confidence dengan support yang dipilih sebelumnya. Dalam menemukan support yang sesuai peneliti melakukan trial dan error dalam mencoba support untuk dimasukkan ke dalam plot dan dipilih yang sesuai setelahnya. Pemilihan support dan confidence dari trial dan error diharapkan dari peneliti nantinya akan menghasilkan rule tidak lebih dari 25 rules dan tidak kurang dari 20 rules. Pada penelitian ini setelah dilakukan beberapa kali pemilihan support, maka akan digunakan support sebesar 0.06. Gambar 5 merupakan plot yang supportnya dipilih 0.06.



Gambar 5. Plot dengan support 0.6

Pada plot tersebut dapat diamati pada support 0.06 bahwa confidence dari 0.8 hingga 1.0 ditemukan rules yang berjumlah 0. Sedangkan pada confidence 0.07 dan 0.06 rules yang ditemukan terlalu sedikit. Sehingga peneliti mengambil nilai confidence yang 0.05. Setelah memilih minimum support dan confidence tersebut kemudian mencari rules menggunakan algoritma apriori dengan bantuan aplikasi GUI R.

Rules yang dibangkitkan menggunakan algoritma apriori pada transaksi bulan Mei hingga Oktober ditampilkan pada tabel yang ada pada tabel 1 yang berisi rules dengan nilai ukuran kebaikannya. Hasil tersebut didapatkan dari pencarian frequent itemset dengan nilai minimum

support 0.06 dan confidence 0.5 pada seluruh transaksi dan terdapat 22 rules yang dihasilkan. Perhitungan dari hasil diperoleh menggunakan gui R pada menu analisis keranjang belanja yang sintaksnya terdapat pada lampiran 4. Pada 22 rules yang telah dihasilkan nilai support tertinggi dimiliki oleh rules no 1, confidence tertinggi dimiliki rules no 17, dan lift tertinggi dimiliki oleh rules no 22. Dari hasil rules dapat dibaca jika konsumen membeli produk pada lhs maka konsumen membeli produk pada rhs dengan nilai support dan confidencenya. Nilai support pada rules jika semakin besar menunjukkan bahwa produk yang dibeli bersamaan tersebut sering dibeli dalam transaksi, sebaliknya semakin kecil nilai support menunjukkan produk yang dibeli bersamaan tersebut jarang dibeli dalam transaksi. Sedangkan untuk confidence yang semakin besar maka menunjukkan produk pada rhs sering dibeli bersamaan dengan produk pada lhs.

Dari hasil rules dapat dibaca jika konsumen membeli produk pada lhs maka konsumen membeli produk pada rhs dengan nilai support dan confidencenya. Nilai support pada rules jika semakin besar menunjukkan bahwa produk yang dibeli bersamaan tersebut sering dibeli dalam transaksi, sebaliknya semakin kecil nilai support menunjukkan produk yang dibeli bersamaan tersebut jarang dibeli dalam transaksi. Sedangkan untuk confidence yang semakin besar maka menunjukkan produk pada rhs sering dibeli bersamaan dengan produk pada lhs. Berikut merupakan hasil analisis dengan nilai minimum support 0.06 dan minimum confidence 0.5 yang menghasilkan 22 rules:

Tabel 1. Hasil pembangkitan rules

No	LHS	RHS	Support	Confidence	Lift
1	{Daun Bawang}	{Wortel}	0.176	0.607	1.585
2	{Tomat Merah}	{Wortel}	0.146	0.617	1.611
3	{Bayam}	{Wortel}	0.138	0.548	1.431
4	{Kentang}	{Wortel}	0.130	0.634	1.655
5	{Brokoli}	{Wortel}	0.118	0.594	1.550
6	{Bawang Merah Sedang}	{Wortel}	0.110	0.520	1.358
7	{Kangkung}	{Wortel}	0.105	0.532	1.390
8	{kubis}	{Wortel}	0.105	0.662	1.729
9	{Tauge Panjang}	{Wortel}	0.104	0.534	1.393
10	{Sawi Sendok /Pakcoy}	{Wortel}	0.099	0.544	1.421
11	{Jagung Manis}	{Wortel}	0.097	0.564	1.473
12	{Buncis}	{Wortel}	0.092	0.642	1.676
13	{Labu siam}	{Wortel}	0.086	0.638	1.666
14	{kubis}	{Daun Bawang}	0.083	0.527	1.820
15	{Bawang Putih}	{Bawang Merah Sedang}	0.083	0.700	3.309
16	{Sawi Putih}	{Wortel}	0.079	0.596	1.555

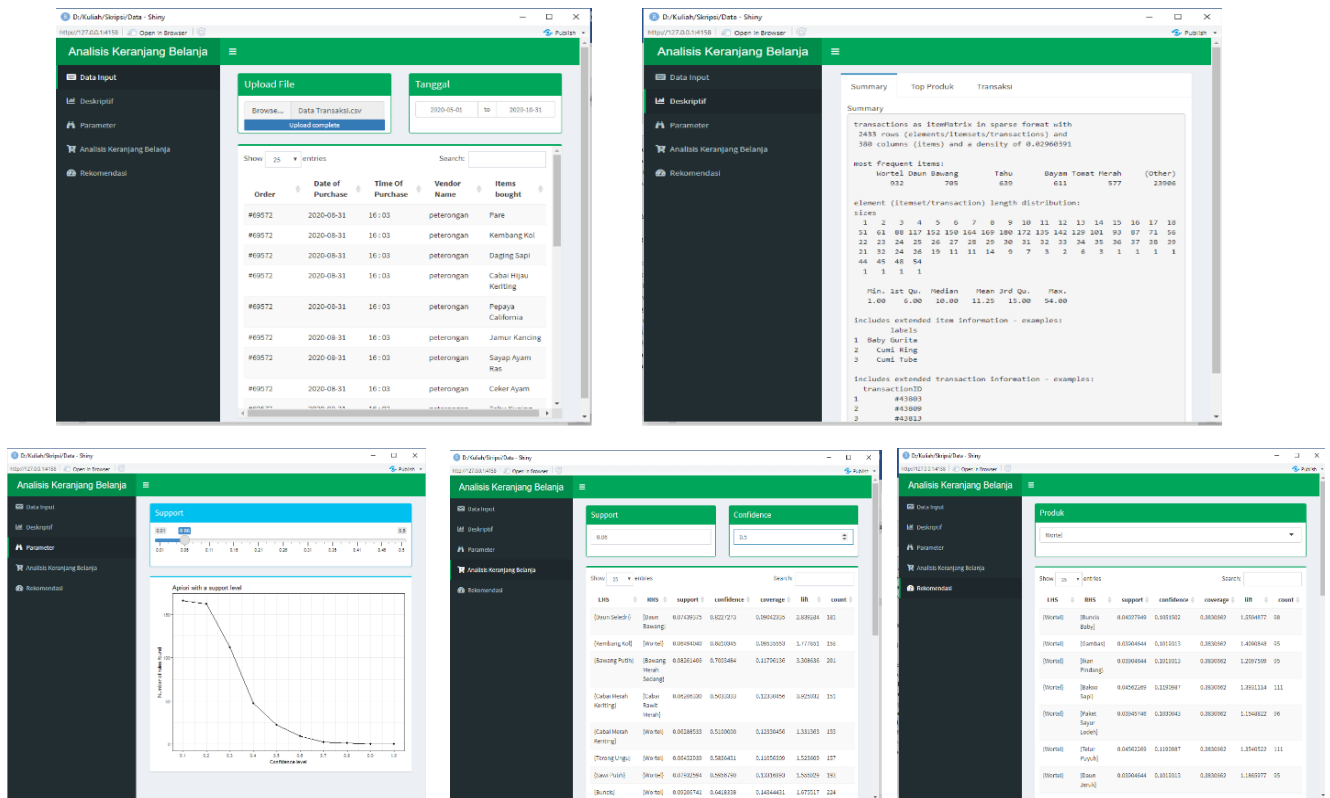
Tabel 2. Hasil pembangkitan rules (Lanjutan)

17	{Daun Seledri}	{Daun Bawang}	0.074	0.823	2.839
----	----------------	---------------	-------	-------	-------

18	{Labu siam}	{Daun Bawang}	0.072	0.529	1.825
19	{Kembang Kol}	{Wortel}	0.065	0.681	1.778
20	{Terong Ungu}	{Wortel}	0.065	0.584	1.524
21	{Cabai Merah Keriting}	{Wortel}	0.063	0.510	1.331
22	{Cabai Merah Keriting}	{Cabai Rawit Merah}	0.062	0.503	3.925

GUI yang disusun pada penelitian ini menggunakan pemrograman bahasa R yang dibuat menggunakan software R studio. Pembuatannya dilakukan dengan bantuan paket shiny dan shiny dashboard sebagai paket utama dalam pembuatan GUI serta beberapa paket lainnya sebagai pendukung pengolahan dan visualisasi data. Tampilan GUI R ini terdiri dari 3 bagian, meliputi header, Sidebar Menu, dan body dashboard. Pada sidebar menu terdapat 5 menu, yaitu menu “data input”, menu “deskriptif”, menu “parameter”, menu “analisis keranjang belanja”, dan menu “rekomendasi”.

GUI R ini digunakan untuk melakukan analisis keranjang belanja pada data transaksi yang diperlukan hanya menginput data dan menentukan nilai support dan confidence sesuai yang diinginkan. Pada menu “input data” digunakan untuk mengunggah dan menampilkan data. Pada menu “deskriptif” terdapat 3 sub menu yang menunjukkan ringkasan data transaksi, data top produk, dan juga data banyaknya transaksi yang terjadi berdasarkan bulan, hari, dan jam. Pada menu “parameter” dapat melihat seberapa banyak rule yang akan dihasilkan pada support tertentu yang dapat diatur dengan menggeser slider yang nantinya akan diperlihatkan dengan sebuah plot dengan nilai x-axis nya adalah confidence dan y-axisnya adalah banyaknya rule yang akan dihasilkan. Dari plot yang dihasilkan dapat dipilih support dan confidence yang ingin digunakan sesuai banyaknya aturan yang dihasilkan.

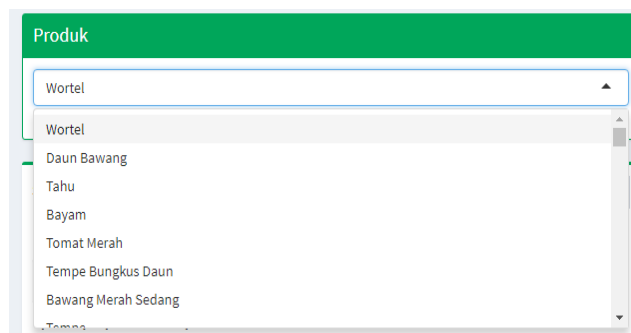


Gambar 6. Ilustrasi penggunaan GUI Analisis Keranjang Belanja

Setelah memilih support dan confidence, maka pada menu “analisis keranjang belanja” dapat diisi support dan confidence yang akan menghasilkan tabel yang berisi aturan-aturan yang sesuai dengan batas support dan confidencenya. Pada menu “rekomendasi” didalamnya terdapat produk-produk yang ada pada data yang setelah dipilih akan menampilkan sebuah tabel yang isinya merupakan produk yang dipilih dan pasangan produk dengan ukuran support dan confidence sesuai dengan menu analisis keranjang belanja.

Pada penelitian ini sistem rekomendasi hanyalah sebagai suatu simulasi yang mengatakan bahwa dengan analisis keranjang belanja dapat juga untuk dibuat sebagai rekomendasi produk. Terlebih pada e-commerce yang memasarkan produknya secara digital. Rekomendasi dapat digunakan menjadi salah satu pilihan untuk meningkatkan penjualan pada perusahaan e-commerce. Seperti pada aplikasi penjualan perusahaan pada tampilan produk yang dipilih konsumen dibawahnya dapat ditampilkan produk lainnya dengan menggunakan data hasil dari analisis keranjang belanja.

Pada GUI R terdapat menu rekomendasi yang di dalamnya dapat memilih sebuah produk dan akan ditampilkan tabel yang berisi produk pilihan dan pasangan produk yang



Gambar 7. Produk pada rekomendasi

sesuai dengan nilai support dan confidencenya. Untuk mendapatkan hasil yang lebih baik pada menu rekomendasi disarankan untuk merubah support dan confidence menjadi nilai yang kecil dikarenakan tidak setiap produk memiliki nilai support dan confidence yang besar bahkan terdapat produk yang tidak menghasilkan rekomendasi karena terlalu sedikit orang yang membeli sehingga hasil support dan confidencenya sangat kecil. Oleh karena itu, pembangkitan rules untuk menu rekomendasi menggunakan nilai support 0.0001 dan confidence 0.1.

Pada gambar 7 merupakan beberapa produk yang dapat dipilih pada menu rekomendasi. Contoh bila dipilih produk wortel maka akan menghasilkan tabel pada lampiran 3. Dari tabel tersebut terdapat 42 rules yang dihasilkan dimana pada sisi kiri atau lhs merupakan produk wortel semua sedangkan pada sisi kanan atau rhs berisi produk bermacam jenis lainnya dan merupakan pasangan dari wortel dengan support dan confidencenya masing-masing. Hasil pasangan dari wortel didapatkan dari histori transaksi selama 6 bulan dan dari beberapa hasil tersebut dapat direkomendasikan kepada pembeli yang membeli wortel. Untuk memilih produk mana saja yang akan ditampilkan untuk direkomendasikan dapat dinilai dari ukuran nilai support dan confidence yang bernilai cukup besar.

5. KESIMPULAN

Berdasarkan hasil dan pembahasan pada bab sebelumnya diperoleh kesimpulan sebagai berikut :

1. Penerapan algoritma apriori untuk analisis keranjang belanja menghasilkan sebanyak 22 rules dengan menggunakan minimum support 0.06 dan confidence 0.5. Semakin besar nilai support maka semakin sering produk atau rule dibeli oleh konsumen dari seluruh data transaksi dan sebaliknya. Sedangkan semakin besar nilai confidence maka semakin sering rule tersebut dibeli bersamaan.
2. Analisis keranjang belanja menghasilkan analisis berupa kebiasaan pelanggan dalam membeli produk secara bersamaan. Dari informasi tersebut dapat digunakan untuk melakukan promosi untuk meningkatkan penjualan oleh perusahaan. Contoh yang dapat dilakukan adalah memberikan diskon suatu produk jika membeli produk lainnya.
3. Hasil analisis keranjang belanja dilakukan dengan menggunakan GUI R yang telah dibuat. GUI tersebut memudahkan pengguna untuk melakukan analisis keranjang belanja dengan algoritma apriori. Selain itu, GUI juga ditambahkan menu rekomendasi sebagai contoh penggunaan analisis keranjang belanja jika ingin dijadikan sistem rekomendasi walaupun memiliki banyak keterbatasan.

DAFTAR PUSTAKA

- Agrawal, C. C. 2016. *Recommender Systems: The Textbook*. New York: Springer.
- Berry, M. W., Browne, M. 2006. *Lecture Notes in Data Mining*. Singapore: World Scientific Publishing Co. Pte. Ltd.
- Haryanto, D., Oslan, Y., Dwiyan, D. 2011. *Implementasi Analisis Keranjang Belanja Dengan Aturan Asosiasi Menggunakan Algoritma Apriori Pada Penjualan Suku Cadang Motor*. Jurnal Buana Informatika Vol. 2, No. 2 : Hal 81-94.
- Hermawati, F. A. 2013. *Data Mining*. Yogyakarta: ANDI.
- Kurniawan, F. et al. 2018. *Market Basket Analysis to Identify Customer Behaviours by Way of Transaction Data*. Knowledge Engineering and Data Science Vol. 1, No. 1 : Hal 20-25.
- Limbong, T. 2019. *E-Commerce: Implementasi, Strategi dan Inovasinya*. Medan: Yayasan Kita Menulis.
- Linoff, G. S., Berry, M. J. A. 2011. *For Marketing, Sales, and Customer Relationship Management*. Canada: Wiley Publishing, Inc.
- Listriani, D., Setyaningrum, A. H., Eka, F. 2016. *Penerapan Metode Asosiasi Menggunakan Algoritma Apriori pada Aplikasi Analisa Pola Belanja Konsumen (Studi Kasus Toko Buku Gramedia Bintaro)*. Jurnal Teknik Informatika Vol. 9, No. 2 : Hal 120–127.
- Mundir. 2012. *Statistik Pendidikan*. Jember: STAIN Jember Press.
- Prasetya, C. S. D. 2017. *Sistem Rekomendasi Pada E-Commerce Menggunakan K-Nearest Neighbor*. Jurnal Teknologi Informasi dan Ilmu Komputer Vol. 4, No 3: Hal 194-200.
- Prasetyo, N. A., Alhaq, A. M. G., Bramanto, T. 2020. *Jurus Praktis Statistik Dasar dan Penggunaan R Commander*. Banyumas: CV. Pena Persada
- Ruth, Z. 2003. *E-Commerce and E-Business*. Manila : E-ASEAN Task Force ; Kuala Lumpur : UNDP-APDIP.
- Santoso, P. B. and Hamdani, M. 2007. *Stasistika Deskriptif dalam Bidang Ekonomi dan Niaga*. Jakarta: Erlangga.
- Soliman, F., Youssef, M. A. 2003. *Internet-based e-commerce and its impact on manufacturing and business operations*. Industrial Management and Data Systems Vol. 103, No. 8–9 : Hal 546–552.

- Susanto, S., Suryadi, D. 2010. *Pengantar Data Mining*. Yogyakarta: ANDI.
- Tampubolon, K., Saragih, H., Reza, B. 2013. *Implementasi Data Mining Algoritma Apriori pada Sistem Persediaan Alat-Alat Kesehatan*. *Informasi dan Teknologi Ilmiah* Vol. 1, No. 1 : Hal 93-106.
- Tan, P.-N., Steinbach, M., Kumar, V. 2006. *Introduction to Data Mining*. Boston: Pearson Education.
- Tirta, I. M. 2015. *Panduan Penyusunan Modul dan Analisis Data Online Berbasis Web Interaktif menggunakan R dalam Rangka Melengkapi K Statistika Virtual (Online, Interaktif, Terintegrasi)*. Jember: Universitas Jember.