

Aplikasi Data Mining Asosiasi Barang Menggunakan Algoritma Apriori-TID

I Made Dwi Putra Asana*, I Gede Iwan Sudipa**, A.A. Tri Wulandari Mayun***, Ni Putu Suci Meinarni****, Devi Valentino Waas*****

Teknik Informatika, Fakultas Teknologi dan Informatika, Institut Bisnis dan Teknologi Indonesia

*dwiputraasana@instiki.ac.id, **iwansudipa@instiki.ac.id, ***gungwulan@gmail.com, ****sucimeinarni@instiki.ac.id, ***** valentinowaas30@instiki.ac.id

ABSTRACT

Information about consumer spending patterns is needed by retail companies in determining promotion strategies, placing merchandise on shelves, and inventory management. Most applications in retail companies only produce information from the company side. Applications owned by the company have not been able to generate transaction activity information from the customer side. The transaction activity information from the customer side is knowing consumer spending patterns in terms of buying goods. This information can overcome problems that are often faced by companies, namely customer complaints when there is a vacancy in inventory of pairs of goods that are often purchased by customers. This article describes the application of association data mining with the Apriori TID algorithm. Apriori-TID is a development of the Apriori algorithm. Apriori-TID plays a role in finding frequent itemsets based on *minimum support*. The application can generate goods association rules based on the *minimum support* and confidence values entered by the user. The association data mining application features dataset import, dataset analysis, and print analysis result reports. The association rules generated by the application are evaluated using the lift ratio method. The number of transactions, *minimum support*, and *minimum confidence* affect the number of rules generated by the application.

Keyword: Data Mining, Association, Frequent Itemset, Apriori-TID, Retail

1. Pendahuluan

Pencarian pola belanja konsumen menjadi tantangan bagi perusahaan ritel dalam menentukan strategi penjualan. Manajemen data transaksi berbasis komputer pada umumnya hanya memberikan informasi aktivitas transaksi dari sisi perusahaan. Informasi aktivitas transaksi dari sisi perusahaan yang dihasilkan meliputi rekapitulasi penjualan, pengadaan persediaan, nilai persediaan, dan performa penjualan pegawai. Data transaksi pada basis data dapat digali beberapa informasi tersembunyi berdasarkan aktivitas transaksi dari sisi pelanggan, sehingga pihak pengambil keputusan dapat menentukan strategi untuk meningkatkan penjualan.

Toko modern merupakan salah satu bentuk usaha yang memproduksi data transaksi yang cukup besar pada basis data penjualan. Ayunadi Swalan merupakan salah satu toko modern yang terletak di Kota Denpasar, Provinsi Bali. Sebagai toko modern, Ayunadi Swalayan telah menerapkan teknologi berbasis komputer dalam pencatatan transaksi perusahaan. Pada basis data perusahaan memiliki data item sekitar 15.000 item dan 1.000 transaksi penjualan setiap hari. Data transaksi tercatat sejak tahun 2015 dan terus bertumbuh hingga membentuk data yang cukup besar. Aplikasi yang dimiliki menghasilkan informasi aktivitas transaksi meliputi rekapitulasi penjualan, pembelian, retur, persediaan barang, dan performa penjualan setiap kasir. Aplikasi yang dimiliki oleh perusahaan belum dapat menghasilkan informasi aktivitas transaksi dari sisi pelanggan. Salah satu informasi aktivitas transaksi dari sisi pelanggan adalah mengetahui pola belanja pelanggan dalam hal membeli barang.

Salah satu informasi pola belanja pelanggan adalah kebiasaan pelanggan dalam membeli barang secara bersamaan. Pelanggan memiliki kebiasaan dalam membeli pasangan barang yang selalu dibeli bersamaan. Informasi tersebut dapat mengatasi permasalahan yang sering dihadapi oleh perusahaan yaitu complain pelanggan saat terjadi kekosongan persediaan barang yang sering dibeli secara bersamaan.

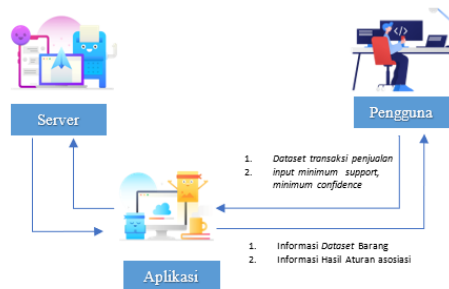
Data transaksi pada basis data yang dimiliki oleh Ayunadi Swalayan yang cukup besar membutuhkan penerapan teknologi dalam menemukan informasi tersembunyi. Teknologi yang dapat diterapkan untuk menemukan pola belanja konsumen dalam membeli barang adalah *data mining*. *Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [1]. Teknik *data mining* yang dapat digunakan untuk mendapatkan pola barang yang sering dibeli secara bersamaan adalah teknik asosiasi. Teknik asosiasi dalam *data mining* adalah menemukan aturan asosiasi yang memenuhi

syarat *minimum support* dan *minimum confidence* [2][3]. Pada teknik asosiasi membutuhkan algoritma pencarian hubungan barang sebelum membentuk aturan asosiasi. Hubungan barang ditemukan dengan algoritma *frequent itemset*.

Algoritma *frequent itemset* yang umum digunakan adalah algoritma Apriori. Penerapan algoritma Apriori pada proses pengolahan data hasil transaksi penjualan di Minimarket Priyo dapat membentuk beberapa pola kombinasi itemsets hasil dan informasi yang dihasilkan dapat digunakan dalam pembuatan katalog produk yang akan dijual [1]. Pada algoritma Apriori untuk menentukan kandidat-kandidat yang mungkin muncul dengan cara memperhatikan *minimum support*, proses utama dalam algoritma Apriori adalah *Join* dan *Prune* [4]. Algoritma Apriori melakukan pemindaian pada database secara berulang kali hingga mendapatkan kandidat sesuai ambang batas yang ditentukan, pemindaian tersebut membutuhkan waktu pemrosesan yang tinggi [5][6]. Masalah pemindaian berulang pada algoritma Apriori diperbaiki pada algoritma Apriori-TID. Salah satu fitur menarik dari algoritma ini adalah bahwa database asli hanya digunakan untuk pemindaian kandidat pasangan barang pada tahap pertama [7]. Apriori-TID memotong dataset setiap level pembentukan itemset. Hal tersebut dapat mengurangi waktu pemrosesan dalam pencarian pasangan barang. Pada artikel ini memaparkan aplikasi *data mining* asosiasi dengan algoritma Apriori-TID dalam pencarian *frequent itemset*. Aplikasi yang dibangun dapat menghasilkan informasi aturan asosiasi barang berdasarkan *minimum support* dan *minimum confidence*. Aturan asosiasi barang yang dihasilkan dievaluasi dengan metode *lift ratio*. Artikel ini terdiri beberapa sesi meliputi pendahuluan, metode penelitian, hasil dan pembahasan, dan kesimpulan.

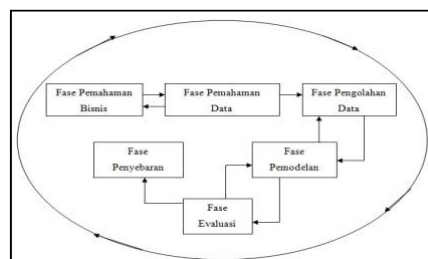
2. Metode Penelitian

Secara umum aplikasi *data mining* asosiasi ditunjukkan pada Gambar 1. Sistem *data mining* asosiasi ini diawali dengan penggalian pola frekuensi tinggi (*frequent pattern mining*) menggunakan Algoritma Apriori-TID sesuai dengan nilai *minimum support* yang dimasukkan oleh pengguna sistem. Setelah ditemukan pola frekuensi tinggi, selanjutnya dilakukan pembentukan aturan asosiasi berdasarkan nilai *minimum confidence* yang dimasukkan oleh pengguna sistem. Kemudian sistem akan memberikan keluaran berupa aturan asosiasi antar item yang dibeli oleh pelanggan



Gambar 1. Gambaran Umum Sistem Aplikasi *Data Mining* Asoosiasi

Pengembangan aplikasi *data mining* asosiasi menggunakan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM). CRISP-DM adalah standar proses *data mining* sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian [8].



Gambar 2. Proses Pengembangan Aplikasi *Data Mining* dengan CRISP-DM

Penerapan CRISP-DM dalam pengembangan aplikasi data mining asosiasi seperti ditunjukkan pada Gambar 2, meliputi fase sebagai berikut :

1. Fase Pemahaman Bisnis

Menentukan tujuan dan kebutuhan secara detail dalam lingkup bisnis secara keseluruhan. Menyiapkan strategi awal untuk mencapai tujuan. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan *data mining*.

2. Fase Pemahaman Data

Mencari pengetahuan awal dan mengenali data lebih jauh dengan cara melakukan analisis penyelidikan data. Dalam fase ini dilakukan eksplorasi data terhadap basis data aplikasi POS Ayunadi Swalayan.

3. Fase Pengolahan Data

Memilih kasus dan variabel yang ingin dianalisis. Pada fase ini dilakukan transformasi data transaksi penjualan berupa nomer transaksi (*transaction id*), kode barang, dan nama barang. Hasil transformasi disimpan pada data aplikasi *data mining* asosiasi.

4. Fase Pemodelan

Pada fase pemodelan, metode yang digunakan untuk menemukan pasangan barang yang sering muncul (*frequent itemset*) adalah Apriori-TID. Penemuan pasangan barang ditentukan berdasarkan *minimum support* yang ditentukan oleh pengguna. Hasil *frequent itemset* kemudian membentuk aturan asosiasi berdasarkan nilai *minimum confidence*.

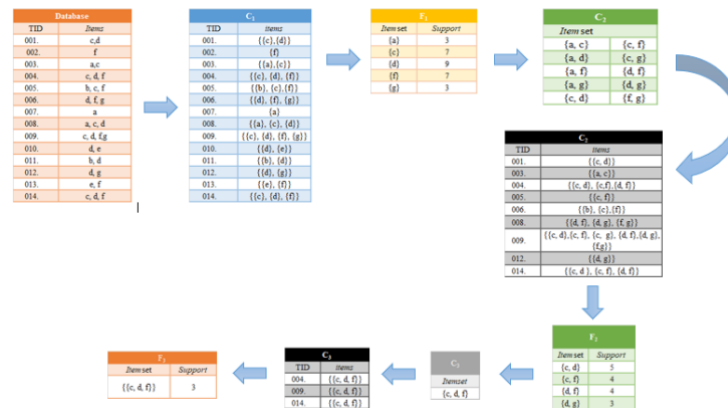
5. Fase Evaluasi

Pada fase ini dilakukan evaluasi terhadap pemodelan yang digunakan. Metode evaluasi yang digunakan adalah lift ratio. Lift ratio mengukur kekuatan aturan asosiasi yang dihasilkan pada fase pemodelan.

6. Fase Penyebaran

Pada fase ini aplikasi *data mining* asosiasi dipasang pada computer perusahaan. Pada fase penyebaran dilakukan pembuatan laporan mengenai informasi kebiasaan belanja konsumen yang telah dihasilkan pada saat pemodelan.

2.1. Apriori TID



Gambar 3. Proses Algoritma Apriori-TID

Gambar 3 menunjukkan tahapan dalam menemukan *frequent itemset* pada Algoritma Apriori-TID. Algoritma Apriori-TID adalah salah satu bagian dari *association rule mining*. *Association rule mining* merupakan salah satu teknik yang ada dalam penambangan data yang bertujuan untuk mendapatkan aturan asosiasi atau relasi antara sekumpulan item. Seperti Algoritma Apriori pada umumnya, Algoritma Apriori-TID juga menggunakan fungsi *Apriori-gen* dalam menentukan kandidat itemset, tetapi perbedaannya terletak pada database yang tidak digunakan dalam perhitungan support setelah lintasan pertama [9][10].

2.2. Lift Ratio

Lift ratio adalah alat ukur penting dalam aturan asosiasi. Fungsinya adalah mengukur ketepatan dan kecermatan suatu alat ukur (*support* dan *confidence*) agar dapat dipercaya sepenuhnya. Dalam artikel ini lift ratio memastikan bahwa apakah barang A dibeli secara bersamaan dengan barang B. Sebuah kombinasi item set dinyatakan valid dan kuat jika nilai *lift ratio* > 1 [11]. Untuk mendapatkan nilai *lift ratio* diawali dengan menghitung nilai *benchmark confidence* yang ditunjukkan pada persamaan 1 sebagai berikut :

$$Benchmark\ Confidence = \frac{\sum Transaksi\ mengandung\ B}{\sum Transaksi} \times 100 \tag{1}$$

Perhitungan nilai *lift ratio* pada sebuah aturan asosiasi dilakukan menggunakan persamaan 2 sebagai berikut :

$$Lift\ Ratio = \frac{Benchmark\ Confidence}{Confidence} \tag{2}$$

3. Hasil dan Analisis

Hasil dari pengembangan aplikasi *data mining* menghasilkan sebuah aplikasi *data mining* berbasis desktop yang dibangun dengan Bahasa pemrograman JAVA. Aplikasi *data mining* yang dihasilkan memberikan informasi mengenai asosiasi barang berdasarkan data transaksi penjualan pada basis data perusahaan. Aplikasi *data mining* menerapkan algoritma Apriori-TID dalam proses pencarian *frequent itemset*. Pada bagian ini menjelaskan menu aplikasi *data mining* asosiasi, pembahasan hasil pencarian aturan asosiasi, pengujian lift ratio, dan pembahasan algoritma Apriori-TID.

3.1. Halaman Beranda

Gambar 4 Merupakan Tampilan Halaman Beranda yang merupakan tampilan awal saat membuka aplikasi. Halaman ini menampilkan Logo dari Ayunadi Swalayan, dan menampilkan empat menu yaitu halaman dataset transaksi penjualan, halaman analisis dataset transaksi penjualan, dan halaman bantuan.



Gambar 4. Halaman Beranda Aplikasi *Data mining* Asosiasi

3.2. Halaman Dataset Transaksi Penjualan

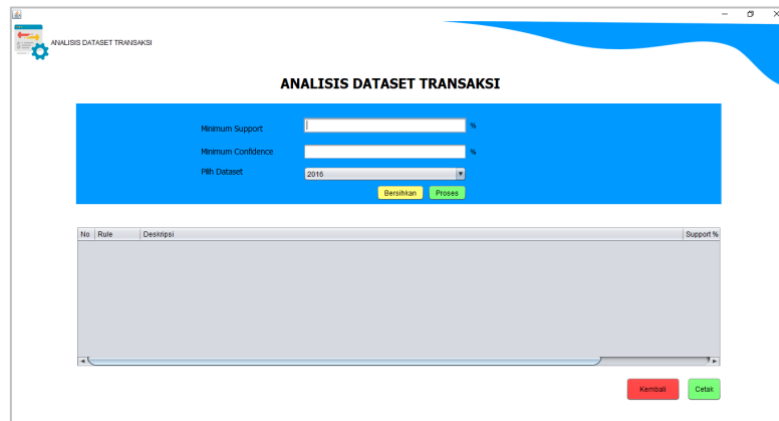
Gambar 5 Merupakan Tampilan Halaman Dataset transaksi penjualan halaman ini digunakan untuk *import* dataset transaksi penjualan, dan dapat melihat dataset barang yang telah di impor. Pengguna dapat menambahkan dataset dengan menekan tombol Impor. Dataset transaksi yang telah diimpor ditunjukkan pada Gambar 5.



Gambar 5. Halaman Dataset Transaksi Penjualan

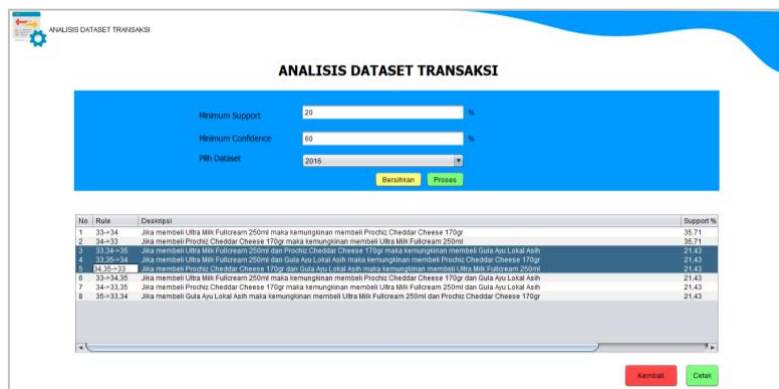
3.3. Halaman Analisis Asosiasi Barang

Gambar 6 merupakan Tampilan Halaman Analisis Dataset transaksi penjualan, dimana halaman ini digunakan untuk menganalisis dataset transaksi penjualan menggunakan algoritma apriori-TID. Pada halaman terdapat fitur untuk memasukkan *minimum support* dan *minimum confidence* yang dibutuhkan saat memproses algoritma tersebut dan terdapat tombol simpan untuk menyimpan data hasil analisis. Gambar 6 menunjukkan kondisi awal dari halaman analisis dataset transaksi.



Gambar 6. Halaman Analisis Dataset Transaksi

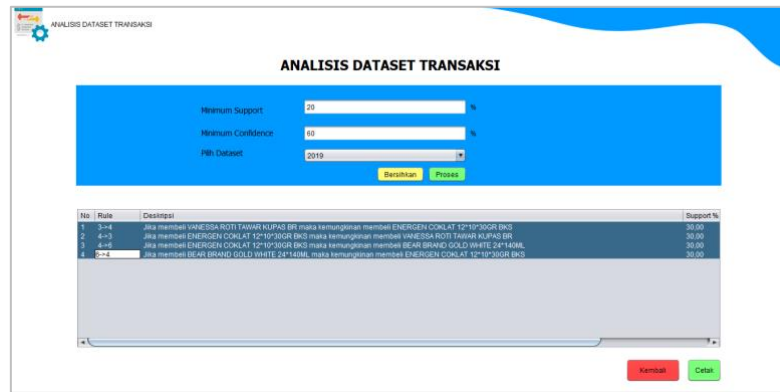
Gambar 7 menunjukkan hasil analisis dataset transaksi berupa aturan hubungan barang. *Minimum support* yang digunakan adalah 20% dan *minimum confidence* sebesar 60%. Hasil aturan asosiasi dapat dilihat pada Gambar 7.



Gambar 7. Hasil Eksperimen Analisis Asosiasi Pertama

Pada Gambar 7, data transaksi yang menjadi ujicoba adalah sebanyak 14 transaksi. Hasil eksperimen dari 14 data transaksi dengan *minimum support* 20% dan *minimum confidence* 60% menghasilkan 8 jumlah aturan, namun yang memenuhi syarat lift ratio >1 hanya 3 rule yaitu:

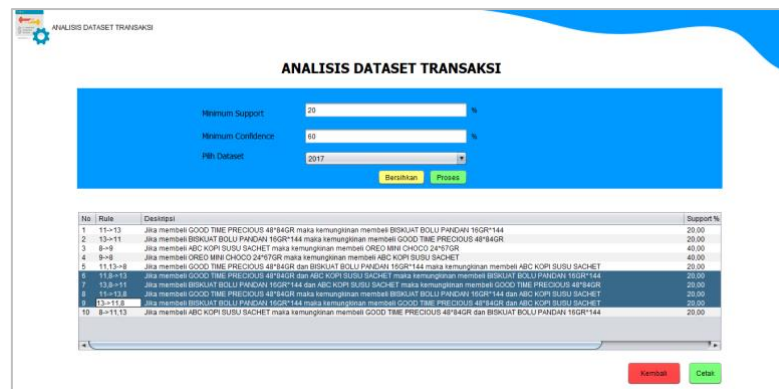
- 1) Jika membeli Ultra Milk Fullcream 250ml dan Prochiz Cheddar Cheese 170gr maka kemungkinan membeli Gula Ayu Lokal Asih (support : 21,43 , confidance : 60,00 dan lift ratio : 1,20)
- 2) Jika membeli Ultra Milk Fullcream 250ml dan Gula Ayu Lokal Asih maka kemungkinan membeli Prochiz Cheddar Cheese 170gr (support: 21,43 , confidance: 75,00 dan lift ratio : 1,17)
- 3) Jika membeli Prochiz Cheddar Cheese 170gr dan Gula Ayu Lokal Asih maka kemungkinan membeli Ultra Milk Fullcream 250ml (support: 21,43 , confidance :75,00 dan lift ratio :1,50)



Gambar 8. Hasil Eksperimen Analisis Asosiasi Kedua

Gambar 8 merupakan hasil eksperimen dari data transaksi yang berjumlah 50 data. Hasil eksperimen dari 50 data transaksi dengan *minimum support* 20% dan *minimum confidence* 60% menghasilkan 4 jumlah aturan, yang memenuhi syarat lift ratio >1 adalah 4 aturan yaitu:

- 1) Jika Membeli Vanessa Roti Tawar Kupas Br Maka Kemungkinan Membeli Energen Coklat 12*10*30gr Bks
(support : 30,00 , confidance : 65,22 dan lift ratio : 1,02)
- 2) Jika Membeli Energen Coklat 12*10*30gr Bks Maka Kemungkinan Membeli Vanessa Roti Tawar Kupas Br
(support: 30,00, confidance: 46,88 dan lift ratio : 1,02)
- 3) Jika Membeli Energen Coklat 12*10*30gr Bks Maka Kemungkinan Membeli Bear Brand Gold White 24*140ml
(support: 30,00 , confidance :46,88 dan lift ratio :1,02)
- 4) Jika Membeli Bear Brand Gold White 24*140ml Maka Kemungkinan Membeli Energen Coklat 12*10*30gr Bks
(support : 30,00 , confidance : 65,22 dan lift ratio : 1,02)



Gambar 9. Hasil Eksperimen Analisis Asosiasi

Gambar 9 merupakan hasil eksperimen dari data transaksi yang berjumlah 10 data transaksi. Hasil eksperimen dari 10 data transaksi dengan *minimum support* 20% dan *minimum confidence* 60% menghasilkan 10 jumlah aturan, namun yang memenuhi syarat lift ratio >1 hanya 4 aturan yaitu:

- 1) Jika Membeli Good Time Precious 48*84gr Maka Kemungkinan Membeli Biskuat Bolu Pandan 16gr*144
(Support : 20,00 , Confidance : 40,00 Dan Lift Ratio : 2,00)
- 2) Jika Membeli Good Time Precious 48*84gr Dan Biskuat Bolu Pandan 16gr*144 Maka Kemungkinan Membeli Abc Kopi Susu Sachet
(Support : 20,00 , Confidance : 100,00 Dan Lift Ratio : 1,67)
- 3) Jika Membeli Good Time Precious 48*84gr Dan Abc Kopi Susu Sachet Maka Kemungkinan Membeli Biskuat Bolu Pandan 16gr*144
(Support : 20,00 , Confidance : 66,67 Dan Lift Ratio : 3,33)
- 4) Jika Membeli Good Time Precious 48*84gr Maka Kemungkinan Membeli Biskuat Bolu Pandan 16gr*144 Dan Abc Kopi Susu Sachet
(Support : 20,00 , Confidance : 100,00 Dan Lift Ratio : 2,00)

No	Rule	Rule Deskripsi	Support	Confidence
1	3->4	Jika membeli VANESSA ROTI TAWAR KUPAS BR maka kemungkinan membeli ENERGEN COKLAT 12*10*30GR BKS	35.714285	71.42857142
2	4->3	Jika membeli ENERGEN COKLAT 12*10*30GR BKS maka kemungkinan membeli VANESSA ROTI TAWAR KUPAS BR	35.714285	55.55555555
3	3,4->5	Jika membeli VANESSA ROTI TAWAR KUPAS BR dan ENERGEN COKLAT 12*10*30GR BKS maka kemungkinan membeli BEAR BRAND GOLD WHITE 24*140ML	21.428571	60.0
4	3,5->4	Jika membeli VANESSA ROTI TAWAR KUPAS BR dan BEAR BRAND	21.428571	75.0

Gambar 10. Hasil Cetak Aturan Asosiasi Barang

Hasil aturan asosiasi pada aplikasi dapat dicetak untuk kebutuhan laporan pada perusahaan dalam penentuan strategi peletakan barang pada rak barang, pemesanan barang, dan promosi paket barang. Gambar 10 merupakan tampilan cetak aturan asosiasi barang. Tombol cetak dapat diakses pada halaman analisis dataset transaksi (Gambar 6).

4. Kesimpulan

Berdasarkan penelitian pengembangan aplikasi *data mining* asosiasi yang telah dilakukan, dapat disimpulkan bahwa Algoritma Apriori-TID dapat diterapkan dalam proses penemuan *frequent itemset* pada analisis asosiasi. Apriori-TID memangkas dataset penelusuran pasangan barang dimulai dari pencarian 2 item barang berpasangan. Pengembangan aplikasi *data mining* dengan CRISP-DM meliputi fase pemahaman bisnis, fase pemahaman data, fase pengolahan data, fase pemodelan, fase evaluasi, dan fase penyebaran. Aplikasi *data mining* yang dibangun dapat menghasilkan aturan asosiasi berdasarkan *minimum support* dan *minimum confidence* yang ditentukan oleh pengguna sesuai kebijakan perusahaan. Jumlah transaksi, *minimum support*, dan *minimum confidence* mempengaruhi jumlah aturan yang dihasilkan oleh aplikasi. Saran pengembangan aplikasi *data mining* asosiasi dapat dilakukan dengan menambahkan parameter utilitas barang seperti nilai keuntungan barang dan harga pokok barang. Sehingga dapat diketahui peluang keuntungan aturan asosiasi yang dihasilkan oleh aplikasi.

Acknowledgements

Penelitian aplikasi *data mining* asosiasi didukung dan didanai oleh Lembaga Penelitian dan Pengabdian kepada Masyarakat INSTIKI melalui skema penelitian dosen dan mahasiswa tahun 2021 dengan nomor kontrak 02/04/LPPM/PDM/V/2021.

References

- [1] A. Anto, T. Susilo, P. T. Informatika, S. Selatan, and S. Selatan, "Penerapan Algoritma Apriori pada Pengolahan Data Transaksi Penjualan di Minimarket Priyo Kota Lubuklinggau," vol. 01, no. 03, 2018.
- [2] Junaidi, Y. Novianto, and Jasmir, "Penerapan Metode Asosiasi Data Mining Murni Baru Furniture Jambi," *J. Ilm. Mhs. Tek. Inform.*, vol. 2, no. 4, pp. 269–281, 2020.
- [3] M. H. Prayitno and R. Rasim, "Analisa Penjualan Produk Retail Dengan Metode Data Mining Asosiasi," *J. Kaji. Ilm.*, vol. 18, no. 3, p. 231, Sep. 2018, doi: 10.31599/jki.v18i3.273.
- [4] A. Ag. B. Ariana and I. M. D. P. Asana, "ANALISIS KERANJANG BELANJA DENGAN ALGORITMA APRIORI PADA PERUSAHAAN RETAIL," 2013, pp. 2–4.
- [5] C. Yadav, Shuliang Wang, and M. Kumar, "An approach to improve apriori algorithm based on association rule mining," in *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, Jul. 2013, pp. 1–9, doi: 10.1109/ICCCNT.2013.6726678.
- [6] A. Patil and P. Gupta, "A review on up-growth algorithm using association rule mining," in *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, Jul. 2017, pp. 96–99, doi: 10.1109/ICCMC.2017.8282605.
- [7] H. K. D. Sarma and S. Mishra, "Mining Time Series Data with Apriori Tid Algorithm," in *2016 International Conference on Information Technology (ICIT)*, Dec. 2016, pp. 160–164, doi: 10.1109/ICIT.2016.043.
- [8] W. Y. Ayele, "Adapting CRISP-DM for Idea Mining," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 6, 2020, doi: 10.14569/IJACSA.2020.0110603.
- [9] A. M. Siregar and M. Syahrizal, "Implementasi Algoritma Apriori Tid Untuk Mengetahui Pola Penjualan

- Keramik,” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, 2019, doi: 10.30865/komik.v3i1.1572.
- [10] K. Khurana, “A Comparative Analysis of Association Rules Mining,” vol. 3, no. 5, pp. 3–6, 2013, [Online]. Available: <http://www.ijsrp.org/research-paper-0513/ijsrp-p17133.pdf>.
- [11] D. Fitriati, “Implementasi Data Mining untuk Menentukan Kombinasi Media Promosi Barang Berdasarkan Perilaku Pembelian Pelanggan Menggunakan Algoritma Apriori,” *Pros. Annu. Res. Semin. 2016*, vol. 2, no. 1, pp. 472–480, 2016, doi: 979-587-626-0.