

## PENERAPAN METODE ASSOCIATION RULE MINING DALAM PENGEMBANGAN UMKM DENGAN ALGORITMA FP-GROWTH

Antonius Wahyu Sudrajat<sup>1</sup>, Ermatita<sup>2</sup>

<sup>1,2</sup>AMIK MDP, <sup>2</sup>Universitas Sriwijaya Palembang

e-mail :<sup>1</sup>wahyu.sudrajat@mdp.ac.id,<sup>2</sup>ermatita@ilkom.unsri.ac.id

### ABSTRACT

Data mining is an analysis process to find patterns from large data sets, so that it becomes pattern-shaped information that has meaning in decision support. One of the techniques used to find patterns in data sets is the association rule technique. This paper discusses several algorithms in association rule including a priori algorithm, Market-Basket Analysis, frequent Pattern Growth (FP-Growth), associative-classification Method and Multidimensional Association Rule Mining. In this study, the FP-Growth algorithm is used to determine the development pattern of SMEs by looking at the success factors of SMEs development seen from the geographical location and environmental conditions of MSMEs. The results of this study are expected to find the association rules of these success factors, so that they can then be used as a consideration in making a more effective SMEs development strategy.

**Keywords :** Data Mining, Association rule, FP-Growth, SMEs

### INTISARI

Data mining merupakan proses analisis untuk menemukan pola dari kumpulan data yang besar, sehingga menjadi informasi berbentuk pola yang memiliki arti dalam pendukung keputusan. Salah satu teknik yang digunakan untuk menemukan pola dalam kumpulan data adalah teknik association rule. Makalah ini membahas beberapa algoritma dalam association rule diantaranya adalah algoritma apriori, Market-Basket Analysis, frequent Pattern Growth (FP-Growth), associative-classification Method dan Multidimensional Association Rule Mining. Dalam penelitian ini digunakan algoritma FP-Growth dalam mengetahui pola pengembangan UMKM dengan melihat faktor sukses pengembangan UMKM dilihat dari letak geografis dan kondisi lingkungan UMKM. Hasil dari penelitian ini diharapkan menemukan aturan asosiasi dari factor sukses tersebut, sehingga selanjutnya dapat digunakan sebagai pertimbangan dalam membuat strategi pengembangan UMKM yang lebih efektif.

**Kata kunci :** Data Mining, Association rule, FP-Growth, UMKM

### 1. PENDAHULUAN

UMKM (Usaha Mikro, Kecil, dan Menengah) dapat membantu meningkatkan pertumbuhan ekonomi dan menambah pendapatan daerah, sehingga UMKM perlu dilakukan pengelolaan dan pengembangan. Pengembangan UMKM bukan hanya tanggung jawab pemilik namun juga instansi pemerintah khususnya dinas koperasi dan UMKM. Hasil penelitian menunjukkan bahwa ada empat faktor penentu keberhasilan UMKM, yaitu promosi, branding, minat berwirausaha dan jaringan usaha (Tromol & Pabelan, n.d.). Dari tersebut perlu juga dilihat dari letak geografis dan kondisi lingkungan dari UMKM.

Pengembangan UMKM yang dilakukan oleh pemerintah harus berdasarkan data dan informasi yang tepat, sehingga pengembagnan dapat dilakukan lebih terarah dan sesuai dengan kondisi dilapangan. Permasalahan yang terjadi dalam pengembangan UMKM adalah data atau informasi terkait dengan UMKM tidak hanya dimiliki oleh satu instansi saja, namun dimiliki oleh instansi yang berbeda-beda dengan

kemungkinan data yang beragam juga. Selain itu permasalahan yang dihadapi adalah jumlah data yang harus dioleh terlalu banyak. Akibatnya kesalahan dalam pengembangan UMKM sangat besar terjadi.

Penambangan informasi dari database yang besar telah diprediksi oleh banyak peneliti sebagai studi utama dalam sistem yang beragam. Peneliti di berbagai bidang telah memberikan minat yang besar pada data mining (Sharma & Bhatia, 2016). Keberhasilan penambangan data sangat bergantung pada jumlah energi, pengetahuan dan kreativitas yang dimasukkan oleh desainer ke dalamnya (Liang, 2004).

Beberapa teknik data mining yaitu eksplorasi data, segmentasi pelanggan, Kohonen's Self organising map, aturan asosiasi, dan time series digunakan untuk mengekstrak pengetahuan dan pola untuk keputusan strategis pembuatan dan peramalan (Packianather et al., 2017). Salah satu teknik yang banyak digunakan untuk mengetahui pola hubungan sehingga dapat digunakan dalam penentuan strategi adalah assosiasi. Penambangan aturan asosiasi adalah salah satu bidang paling signifikan dalam penambangan data dan penemuan pengetahuan dalam database (Gangurde & Kumar, 2020). Tujuan dalam aturan asosiasi dapat digunakan untuk menemukan hubungan antara produk yang dibeli oleh pelanggan selama berbelanja, dan untuk menentukan kebiasaan membeli pelanggan sejalan dengan data hubungan ini (Yücenur dkk., 2020). Data Mining menggunakan algoritma apriori dapat menghasilkan pola informasi baru untuk menentukan penjualan pola dalam aturan asosiasi barang dan mengembangkan strategi pemasaran untuk memasarkan produk lain dengan melihat apa saja keunggulan produk yang paling banyak terjual (Hermaliani dkk., 2020). Pada penelitian ini, digunakan algoritma FP-Growth untuk menemukan pola pengembangan UMKM yang lebih efektif.

## 2. METODE PENELITIAN

### 2.1 Metode

Penelitian ini diawali dengan melakukan penelitian awal guna mencari dan mempelajari masalah terkait dengan pengembangan UMKM, yang dilanjutkan dengan penentuan ruang lingkup masalah dan latar belakang masalah. Selanjutnya dilakukan pengumpulan data terkait dengan UMKM. Data set yang diperoleh kemudian dilakukan pengolahan terlebih dahulu untuk melihat item set nilai support dan nilai confidence dengan menggunakan software RapidMiner Studio. Tahapan selanjutnya dilakukan analisis masalah dengan menggunakan teknik data mining yaitu algoritma FT-Growth. Hal ini dilakukan agar tercapainya tujuan oleh peneliti dan dapat digunakan dalam pengembangan UMKM kearah yang lebih baik. Langkah-langkah dalam penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

## 2.2 UMKM

Usaha Kecil dan Menengah merupakan sebuah istilah yang mengacu kepada sebuah usaha kecil yang memiliki kekayaan bersih sebesar Rp 200.000.000 dan jumlah tersebut tidak termasuk tanah dan bangunan tempat usaha. Berdasarkan Kepres RI no. 99 tahun 1998 pengertian Usaha Kecil adalah: “Kegiatan ekonomi rakyat yang berskala kecil dengan bidang usaha yang secara mayoritas merupakan kegiatan usaha kecil dan perlu dilindungi untuk mencegah dari persaingan usaha yang tidak sehat.”

UU No. 9 tahun 1995 menerangkan kriteria usaha kecil adalah sebagai berikut (UU No. 9 Tahun 1995 Tentang Usaha Kecil, 1995):

1. Memiliki kekayaan bersih paling banyak Rp. 200.000.000,- (Dua Ratus Juta Rupiah) tidak termasuk tanah dan bangunan tempat usaha.
2. Memiliki hasil penjualan tahunan paling banyak Rp. 1.000.000.000,- (Satu Miliar Rupiah).
3. Milik Warga Negara Indonesia.
4. Berdiri sendiri, bukan merupakan anak perusahaan atau cabang perusahaan yang tidak dimiliki, dikuasai, atau berafiliasi baik langsung maupun tidak langsung dengan Usaha Menengah atau Usaha Besar.
5. Berbentuk usaha orang perorangan, badan usaha yang tidak berbadan hukum, atau badan usaha yang berbadan hukum, termasuk koperasi.

## 2.3 Data Mining

Data mining adalah proses berulang di mana kemajuan ditentukan oleh penemuan, melalui metode otomatis atau manual. Penambangan data paling berguna dalam skenario analisis eksplorasi di mana tidak ada gagasan yang ditentukan sebelumnya tentang apa yang akan merupakan hasil yang "menarik". Data mining adalah pencarian informasi baru, berharga, dan tidak sepele dalam volume data yang besar (Liang, 2004). Data mining adalah menemukan pola yang berguna dalam data. Data mining juga disebut sebagai penemuan pengetahuan, pembelajaran mesin, dan analitik prediktif (Chauhan & Kaur, 2013).

## 2.4 Association Rule

*Association rule* merupakan salah satu teknik utama dalam data mining dan mungkin bentuk paling umum dalam penemuan pola untuk sistem pembelajaran tanpa pengawasan. Teknik ini mengambil semua kemungkinan pola yang menarik dalam database, yang artinya teknik ini tidak meninggalkan kebutuhan bisnis yang terlewat (Liang, 2004). Dua ukuran parameter dalam analisis asosiasi yaitu *support* dan *confidence*.

Support merupakan pengukuran untuk menunjukkan seberapa besar tingkat dominasi suatu item dari keseluruhan transaksi.

$$Support(A) = \frac{\text{jumlah transaksi mengandung } A}{\text{Total Transaksi}} \quad (1)$$

Sedangkan *confidence* merupakan pengukuran untuk menunjukkan hubungan dua item berdasarkan suatu kondisi tertentu.

$$Confidence = \frac{\sum \text{Transaksi mengandung } A \text{ dan } B}{\sum \text{Transaksi mengandung } A} \quad (2)$$

Beberapa teknik yang digunakan untuk association rule mining diantaranya adalah (Liang, 2004)

**a. Algorithm Apriori**

Algoritma Apriori menghitung kumpulan item yang sering dalam *database* melalui beberapa iterasi. Iterasi *i* menghitung semua *i*-itemsets yang sering (itemsets dengan elemen *i*). Setiap iterasi memiliki dua langkah: pembuatan kandidat serta penghitungan dan pemilihan kandidat (Liang, 2004). Contoh kasus penggunaan metode ini adalah untuk mengetahui hubungan antara demografi karakteristik perusahaan dan fasilitas yang diperlukan untuk situs web komersial. Dengan menggunakan sampel 100 perusahaan Iran teratas yang diperingkat oleh Institut Manajemen Industri Iran. Untuk mengumpulkan data, daftar periksa yang dimodifikasi digunakan yang mencakup tiga area fasilitas dalam situs web komersial, yaitu fasilitas dasar, penyediaan informasi, dan penyampaian layanan (Azimi & Kaffashpour, 2013).

**b. Market-Basket Analysis**

*Market Basket Analysis* adalah metode pemodelan yang didasarkan pada gagasan bahwa ketika konsumen membeli satu set barang tertentu, pembeli memiliki kemungkinan yang lebih tinggi atau lebih rendah untuk membeli satu set produk lain. Dalam menentukan tata letak dan mempromosikan barang-barang di dalam toko, analisis keranjang pasar dapat digunakan. Ketika pembeli *X* lebih cenderung membeli *Y*, seperti yang telah diidentifikasi, margin tinggi *Y* dapat ditemukan di dekat etalase *X*. Konsumen yang telah membeli *Y* akan cukup diyakinkan untuk juga membeli *X* (Sudirman et al., 2020). Teknik atau metode ini juga dikenal sebagai pembelajaran aturan asosiasi atau analisis afinitas, adalah teknik data mining yang dapat digunakan di berbagai bidang, seperti pemasaran, bioinformatika, bidang pendidikan, ilmu nuklir dll (Kaur & Kang, 2016).

**c. Associative-Classification Method**

*Classification based on multiple association rules* (CMAR) adalah metode klasifikasi yang diadopsi dari metode pertumbuhan pola sering (FP-growth) untuk menghasilkan kumpulan item yang sering (Liang, 2004). Metode CMAR terdiri dari dua tahap, yaitu pembuatan atau pelatihan aturan dan Klasifikasi atau pengujian.

**d. Multidimensional Association Rule Mining**

Teknik ini membagi proses penambangan menjadi dua langkah: pertama menambang pola tentang informasi dimensi dan kemudian menemukan kumpulan item yang sering dari sub database yang diproyeksikan, atau sebaliknya (Liang, 2004).

**e. Frequent Pattern Growth (FP-GROWTH) Method**

Frequent pattern growth (FP-growth) adalah cara yang efisien untuk menambang frequent itemsets dalam database yang besar (Liang, 2004). FP-growth adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data (Tana dkk., 2018). Penggalan itemset yang frequent dengan menggunakan algoritma FP-Growth akan dilakukan dengan cara membangkitkan struktur data tree (FP-Tree). Dalam Teknik FP-Growth dapat dibagi menjadi tiga tahapan yaitu (Han dkk., 2012):

1. Tahapan pembangkitan *Conditional Pattern Base*

*Conditional Pattern Base* merupakan subdatabase yang berisi lintasan *prefix* dan *suffix pattern*. Pembangkitan *conditional pattern base* didapatkan melalui FP-Tree yang telah dibangun sebelumnya

2. Tahapan pembangkitan *Conditional FP-Tree*

Support count dari setiap item pada setiap *conditional pattern base* dijumlahkan, lalu setiap item yang memiliki jumlah support count lebih besar sama dengan minimum support count  $\xi$  akan dibangkitkan dengan *conditional FP-Tree*.

### 3. Tahap pencarian *Frequent Itemset*

Apabila Conditional FP-Tree merupakan lintasan tunggal, maka didapatkan *frequent itemset* dengan melakukan kombinasi item untuk setiap *conditional FP-Tree*. Jika bukan lintasan tunggal, maka dilakukan pembangkitan FP-Growth secara rekursif.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Pengolahan Data

Tabel 1 merupakan data UMKM berdasarkan faktor sukses yang telah ditentukan sebelumnya dan diolah menggunakan Microsoft excel.

**Tabel 1.** Data UMKM

ID	Faktor
1	padat bersih, dekat kota, dekat pasar, transportasi mudah
2	padat bersih, dekat tempat wisata, dekat pasar
3	padat kumuh, dekat pasar, transportasi mudah, dekat wisata
4	padat kumuh, dekat kota, dekat tempat wisata
5	luas kumuh, dekat pasar, dekat kota
6	luas kumuh, dekat tempat wisata
7	luas bersih, transportasi mudah, dekat kota
8	luas bersih, dekat tempat wisata, dekat pasar
9	padat bersih, dekat kota
10	padat kumuh, transportasi mudah, dekat pasar

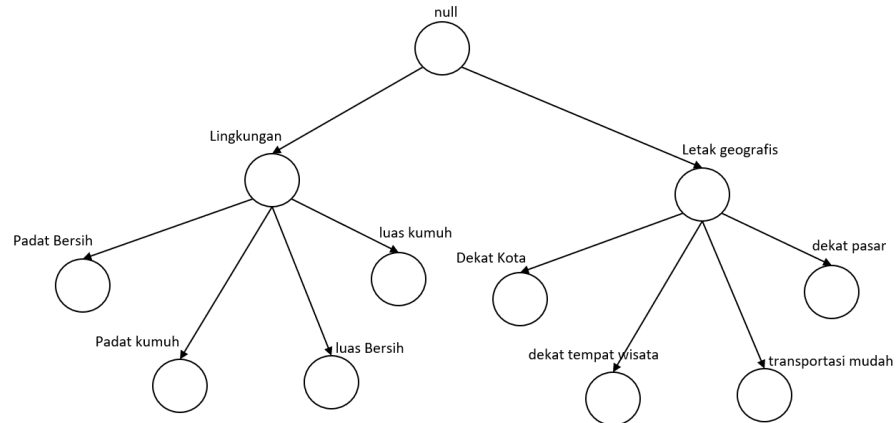
Data UMKM yang diperoleh kemudian dibuat tabular data, yang dikonversi dalam bentuk 1 dan 0 (biner). Data TabularUMKM dapat dilihat pada Tabel 2.

**Tabel 2.** Data Tabular UMKM

ID	Padat Bersih	Padat Kumuh	Luas Bersih	Luas Kumuh	Dekat Pasar	Dekat Tempat wisata	Transportasi Mudah	Dekat Kota
1	1	0	0	0	1	0	1	1
2	1	0	0	0	1	1	0	0
3	0	1	0	0	1	1	1	0
4	0	1	0	0	1	1	0	0
5	0	0	0	1	1	0	0	1
6	0	0	0	1	0	1	0	0
7	0	0	1	0	0	0	1	1
8	0	0	1	0	1	1	0	0
9	1	0	0	0	0	1	0	0
10	0	1	0	0	0	0	1	0

### 3.2 Desain Tree

Berdasarkan data UMKM yang ada selanjutnya di lakukan pembuatan desain Tree. Desain tree dari seluruh data UMKM ditunjukkan pada Gambar 2.



Gambar 2. Pembangkitan FP-Tree

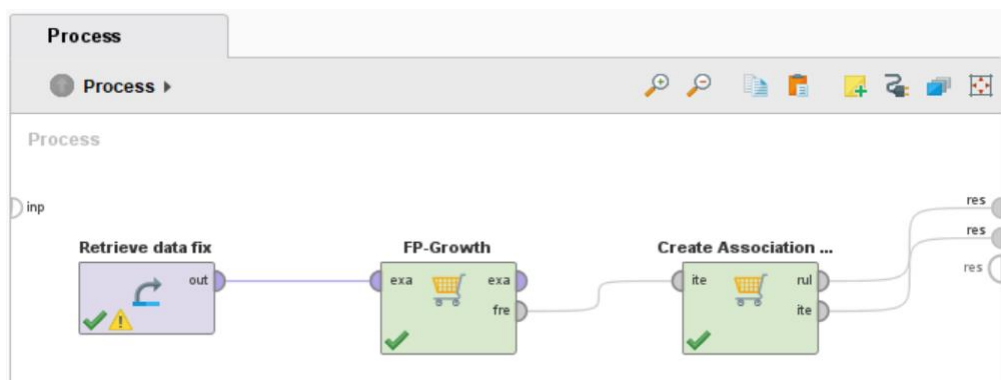
### 3.3 FP Growth

Dalam pengujian dari data UMKM yang telah berupa pola hubungan kombinasi antara items dan association rule sesuai dengan algoritma FP-Growth, maka digunakan aplikasi rapidminer untuk melakukan pengujian terhadap Analisa data UMKM.

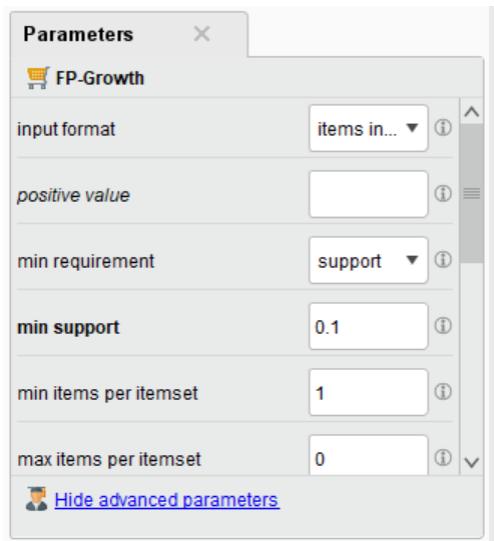
Evaluasi data mining dilakukan dengan menggunakan dataset pengujian yang telah disiapkan, yaitu terdiri data 10 data UMKM. Kemudian dilakukan perhitungan menggunakan algoritma FP-Growth. Dimana langkah-langkah pembentukan model mining menggunakan FP-Growth pada software Rapidminer adalah sebagai berikut:

- Meng-import file format xls dari data UMKM yang telah disiapkan.
- Menerapkan algoritma FP-Growth dalam penentuan asosiasi
- Pengaturan minimum *support*, *confidence* dan *rules* yang dihasilkan.

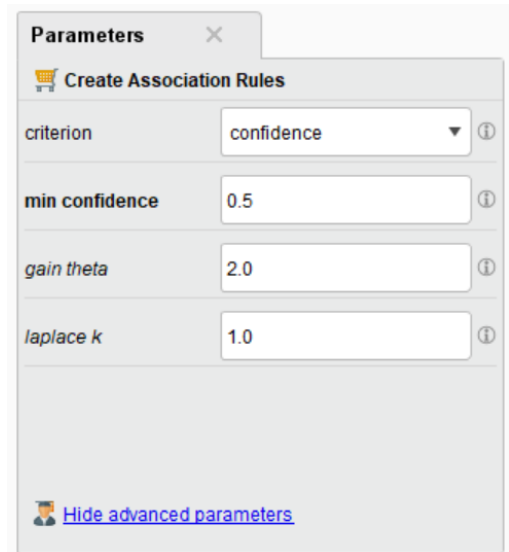
Gambar 3 menunjukkan proses koneksi data dalam software rapidminer dengan menggunakan algoritma FP-Growth. Gambar 4 menunjukkan parameter *support* sedangkan Gambar 5 menunjukkan parameter *confidence*.



Gambar 3. Proses Koneksi data menggunakan FP-Growth



Gambar 4. Parameter Support



Gambar 5. Parameter Confidence

### 3.4 Association Rule

Tahapan ini digunakan untuk menentukan nilai *support* dan *confidence* pada setiap itemset dengan rumus yang sudah dijelaskan sebelumnya. Gambar 6 merupakan hasil dari item set pengolah data dirapidminer.

No.	Premises	Conclusion	Support	Confidence
184	Padat Bersih	Dekat Tempat Wisata	0.417	0.556
185	Padat Bersih	Dekat Kota	0.417	0.556
186	Padat Bersih	Dekat Pasar, Dekat Tempat Wisata	0.417	0.556
187	Padat Bersih	Dekat Pasar, Dekat Kota	0.417	0.556
188	Dekat Tempat Wisata	Dekat Kota	0.333	0.571
189	Dekat Tempat Wisata	Transportasi Mudah	0.333	0.571
190	Dekat Tempat Wisata	Padat Bersih, Dekat Kota	0.333	0.571
191	Dekat Tempat Wisata	Dekat Pasar, Dekat Kota	0.333	0.571
192	Dekat Tempat Wisata	Padat Bersih, Dekat Pasar, Dekat Kota	0.333	0.571
193	Dekat Kota	Padat Kumuh	0.250	0.600
194	Padat Bersih, Dekat Tempat Wisata	Transportasi Mudah	0.250	0.600
195	Dekat Kota	Padat Bersih, Padat Kumuh	0.250	0.600
196	Padat Bersih, Dekat Kota	Padat Kumuh	0.250	0.600

Gambar 6. Hasil dari Item Set

Melihat pada data Gambar 6, maka dapat dilihat bahwa ketika lingkungan UMKM padat bersih dan dekat tempat wisata maka tingkat *confidence* adalah sebesar 50%. Dari item set yang dihasilkan yang memiliki tingkat *confidence* terbesar adalah sebesar 88%, dimana Ketika lingkungan UMKM padat bersih dan letak geografisnya dekat dengan pasar. Gambar 7 adalah hasil dari *association rule* dari data set UMKM.



Gambar 7. Association Rules yang dihasilkan

#### 4. KESIMPULAN

Beberapa kesimpulan yang dapat diambil dari hasil penelitian ini adalah:

1. Data mining menggunakan algoritma FP-Growth dapat menghasilkan pola informasi baru untuk menentukan pola pengembangan UMKM, sehingga dapat memberikan dasar dalam pengembangan UMKM ke arah yang lebih baik.
2. Association rule dapat diterapkan untuk menentukan informasi berupa pola informasi yang dapat digunakan dalam pengembangan UMKM.
3. Besarnya presentase dalam menentukan minimum support dan minimum confidence dipengaruhi oleh data variabel yang digunakan untuk mencari frequent item set yang saling berhubungan untuk menemukan data variabel yang akan dijadikan usulan atau gagasan untuk menganalisis pola pengembangan UMKM.

#### DAFTAR PUSTAKA

- Azimi, A., & Kaffashpour, A. (2013). Applying the Apriori Algorithm for Investigating the Associations Between Demographic Characteristics of Iranian Top 100 Enterprises and the Structure of Their Commercial Websites. *International Journal of Data Mining & Knowledge Management Process*, 3(6), 21–39. <https://doi.org/10.5121/ijdkp.2013.3602>
- Chauhan, R., & Kaur, H. (2013). Predictive analytics and data mining: A framework for optimizing decisions with R tool. In *Advances in Secure Computing, Internet Services, and Applications*. <https://doi.org/10.4018/978-1-4666-4940-8.ch004>
- Gangurde, R. A., & Kumar, B. (2020). Optimal association rule mining for web page prediction using hybrid heuristic trained neural network. *International Journal on Emerging Technologies*, 11(1), 161–173.
- Han, J., Kamber, M., & Pei, J. (2012). Introduction. In *Data Mining*. <https://doi.org/10.1016/b978-0-12-381479-1.00001-0>
- Hermaliani, E. H., Kurniawati, L., Haryanti, T., Mutiah, N., Kurniawan, A., & Renhoran, B. S. (2020). Data



- Mining Technique to Determine the Pattern of Fruits Sales & Supplies Using Apriori Algorithm. *Journal of Physics: Conference Series*, 1641, 012070. <https://doi.org/10.1088/1742-6596/1641/1/012070>
- Kaur, M., & Kang, S. (2016). Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining. *Procedia Computer Science*, 85(Cms), 78–85. <https://doi.org/10.1016/j.procs.2016.05.180>
- Liang, M. (2004). Data Mining: Concepts, Models, Methods, and Algorithms. In *IIE Transactions* (Vol. 36, Issue 5). <https://doi.org/10.1080/07408170490426107>
- Packianather, M. S., Davies, A., Harraden, S., Soman, S., & White, J. (2017). Data Mining Techniques Applied to a Manufacturing SME. *Procedia CIRP*, 62, 123–128. <https://doi.org/10.1016/j.procir.2016.06.120>
- Sharma, S., & Bhatia, S. (2016). Analysis of association rule in data mining. *ACM International Conference Proceeding Series, 04-05-Marc*. <https://doi.org/10.1145/2905055.2905238>
- Sudirman, I. D., Rahmatillah, I., Luckieta, M., & Campus, B. (2020). *Market Basket Analysis Small Medium Enterprise in Bandung Using*. 7(19), 5094–5102.
- Tana, M. P., Marisa, F., & Wijaya, I. D. (2018). Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Pada Toko Oase Menggunakan Algoritma Apriori. *J I M P - Jurnal Informatika Merdeka Pasuruan*, 3(2). <https://doi.org/10.37438/jimp.v3i2.167>
- Tromol, Y., & Pabelan, P. I. (n.d.). *KEY SUCCESS FACTORS OF SMALL AND MEDIUM ENTERPRISES Aris Wijnarko<sup>1</sup> dan Ihwan Susila<sup>2</sup> Abstrak Pendahuluan*. 67–81.
- UU No. 9 tahun 1995 tentang Usaha Kecil, 11 296 (1995).
- Yücenur, G. N., Yaygan, Y., Tevge, H., & DemİR, G. (2020). *Araştırma Makalesi / Research Article Customer Behavior Analysis by Association Rules Mining Birlikteİik Kuralları ile Müşteri Davranışlarının Analizi*. 9(3), 1188–1209.