

PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA FREQUENT PATTERN - GROWTH UNTUK MENENTUKAN POLA PEMBELIAN PRODUK CHEMICALS

Firmansyah, Odi Nurdyawan

Program Studi Manajemen Informatika D3, Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) IKMI Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, 45131, Cirebon, Indonesia
Firmansyah080502@gmail.com

ABSTRAK

Dalam mengembangkan bisnis perusahaan distribusi produk chemicals, penerapan teknologi informasi menjadi sangat penting untuk menjamin kelancaran penjualan produk kebersihan yang ditawarkan. Perusahaan dihadapkan pada tantangan untuk memberikan pelayanan terbaik kepada pembeli. Menerapkan rencana bisnis yang efektif untuk menghindari kerugian. Oleh karena itu, Untuk meningkatkan penjualan produk yang ditawarkan, perusahaan harus mencari strategi yang bagus. Algoritme FP-Growth merupakan salah satu opsi algoritme alternatif yang bisa dipakai untuk mencari pola kombinasi produk. Selain itu, dalam mengontrol dan mengelola persediaan, penting untuk menentukan data yang akurat. Untuk menentukan strategi penjualan produk sabun, salah satu teknik yang dapat digunakan adalah teknik data mining. Dalam hal ini, Algoritma Fp-Growth ialah pengembangan dari algoritma Apriori yang menggunakan konsep pengembangan pohon untuk mencari frequent itemset yang sering dibeli. Pola belanja yang terbentuk berpengaruh besar terhadap penjualan produk. Proses analisis Metode dimulai dengan menjabarkan dan membuat aturan asosiasi untuk beberapa itemset yang sering muncul. Dengan memformulasikan FP-Tree dan mengurangi penggunaan memori, algoritma FP-Growth dapat mempermudah pencarian itemset yang sering muncul. Pola interaksi produk dalam aturan asosiasi dapat membantu perusahaan dalam menetapkan strategi pemasaran dan penjualan yang lebih baik.

Kata Kunci: Transaksi, Data mining, Association rules, Frequent itemset, Fp-Growth

1. PENDAHULUAN

Saat ini persaingan dalam dunia bisnis terutama dalam distribusi produk kebersihan atau chemicals sangatlah sengit. Dawam Lestari Jaya adalah salah satu distributor sabun dan kebersihan di antara banyak pesaing lainnya. Dalam membangun bisnisnya, PT. Dawam Lestari Jaya yang menyediakan produk chemicals tidak dapat menghindari Penggunaan teknologi informasi untuk memfasilitasi penjualan produk Chemicals saat ini menjadi keharusan, mengingat persaingan bisnis yang sengit di bidang tersebut.

Chemicals merupakan kebutuhan penting dalam menjaga kebersihan. Untuk memperoleh chemicals, dapat dibeli dari distributor, perusahaan penyedia, atau tempat penyedia chemicals lainnya. Transaksi pembelian dan penjualan chemicals terjadi setiap harinya di perusahaan distributor. Dalam hal ini, setiap distributor atau perusahaan harus memiliki sistem pengolahan data untuk mencatat transaksi sehingga dapat dimanfaatkan dalam pembuatan laporan. Laporan tersebut akan memberikan informasi berguna seperti produk chemicals yang sering dibeli atau yang paling laris terjual.

“Penerapan data mining untuk menentukan pola pembelian produk.”

Untuk memberikan pelayanan terbaik kepada pelanggan perusahaan, digunakan metode pengolahan data FP-Growth yang

Mengenal set data yang paling mungkin mengalami pengalaman di antara set data. Herasmus

melakukan penelitian dengan menggunakan algoritma FP-Growth untuk menganalisis Sistem Layanan Pelanggan, dan penelitian tersebut menemukan bahwa rata-rata nilai dukungan adalah 80% dan nilai kepercayaan meningkat lebih dari 80%. Sehingga berdasarkan uraian masalah dan penelitian tersebut, Oleh karena itu diusulkan tugas akhir dengan judul Penerapan data mining menggunakan algoritma Fp-growth untuk menentukan pola pembelian produk chemicals. Tujuan dari penyelesaian tugas akhir dengan judul tersebut adalah untuk mengumpulkan informasi dari data penjualan chemicals sehingga manajer perusahaan dapat menggunakan informasi tersebut dalam pengambilan keputusan atau kebijakan strategis. Dari hasil analisis pola transaksi, diharapkan akan dihasilkan informasi

2. TINJAUAN PUSTAKA

2.1. Produk

Produk memegang peran penting dalam dunia bisnis karena dapat memengaruhi kualitas, daya tarik, dan keputusan pembelian konsumen terhadap perusahaan. Oleh karena itu, perusahaan melakukan tahapan-tahapan seperti penelitian pasar, pengembangan, pengujian, dan peluncuran produk untuk memastikan produk sesuai dengan kebutuhan dan harapan pelanggan dan bersaing di pasar. Stanton mendefinisikan produk sebagai "kumpulan atribut fisik yang saling terkait dan dapat diidentifikasi secara nyata."

Produk chemicals adalah cairan kimia yang berfungsi untuk membersihkan berbagai kebutuhan, Chemicals atau bahan kimia memiliki beraneka ragam misalnya sebagai pembersih alat rumah tangga atau pembersih kebutuhan seperti kerak,minyak, noda membandel lainnya, Produk pembersih ini merupakan campuran asam dan surfaktan yang memiliki pH yang relatif aman bagi penggunaannya. Karena sifatnya yang aman, produk ini sering disebut sebagai cleaning chemicals.

2.2. Data mining

Data mining adalah teknik pengolahan data untuk mengidentifikasi informasi yang bermanfaat dari data yang diambil. Metode ini sangat penting di berbagai bidang, termasuk industri, keuangan, cuaca, ilmu, dan teknologi. Di industri, data mining digunakan untuk memperoleh informasi tentang preferensi konsumen, tren pasar, dan permintaan produk. Dengan informasi ini, perusahaan dapat mengoptimalkan strategi pemasaran dan meningkatkan efisiensi operasional. Di bidang keuangan, data mining digunakan untuk menganalisis data keuangan dan memperoleh wawasan yang bermanfaat tentang keadaan ekonomi dan keuangan, termasuk trend investasi, prediksi pasar saham, analisis risiko kredit, dan penilaian kinerja keuangan perusahaan. Dalam bidang cuaca, data mining digunakan untuk menganalisis data cuaca dan memprediksi cuaca di masa depan. Informasi ini dapat digunakan untuk mengambil keputusan strategis, seperti mengoptimalkan rute penerbangan atau mengelola persediaan bahan makanan. [1]

2.3. Proses Data Mining

Menurut Gorunescu [2] proses *data mining* terbagi dalam tiga yaitu:

1. Eksplorasi data melibatkan berbagai kegiatan seperti membersihkan data, mengubah data, mengurangi dimensi data, memilih fitur, dan lain sebagainya.
2. Pemilihan model yang tepat dan valid merupakan tahap penting dalam pembuatan dan pengujian model. Proses tersebut melibatkan seleksi model yang paling sesuai dengan masalah yang sedang dihadapi, yang dilakukan secara kompetitif.
3. Tahap penting dalam penerapan model adalah penggunaan data baru untuk menghasilkan estimasi kasus yang relevan, sehingga dapat diuji apakah model yang dibangun dapat memberikan solusi yang tepat untuk masalah yang dihadapi.

2.4. RapidMiner

RapidMiner adalah sebuah perangkat lunak dengan sumber terbuka (*open source*) yang dapat digunakan untuk melakukan analisis pada data mining, text mining, dan analisis prediksi. Software ini menyediakan sekitar 500 operator data mining yang

mencakup operator untuk *input*, *output*, *data preprocessing*, dan visualisasi. RapidMiner merupakan perangkat lunak yang mandiri untuk analisis data dan dapat digunakan sebagai mesin data mining yang terintegrasi dalam produknya sendiri. Aplikasi ini bermanfaat untuk keperluan bisnis dan komersial, penelitian, pendidikan dan pelatihan, serta pengembangan prototipe dan aplikasi dengan cepat. RapidMiner juga dapat membantu di semua tahapan proses pembelajaran mesin, termasuk persiapan data, visualisasi hasil, validasi, dan pengoptimalan.[3]

2.5. Association Rules

Association rules adalah teknik dalam data mining yang mencari Semua aturan asosiasi dalam database harus memenuhi syarat minimum untuk support (minsup) dan confidence (minconf). Kriteria-kriteria ini digunakan untuk menemukan aturan asosiasi yang menarik dan relevan, dengan memperhatikan batasan yang telah ditentukan. Teknik ini dapat menemukan pola atau hubungan menarik antara item atau atribut pada dataset.[4]

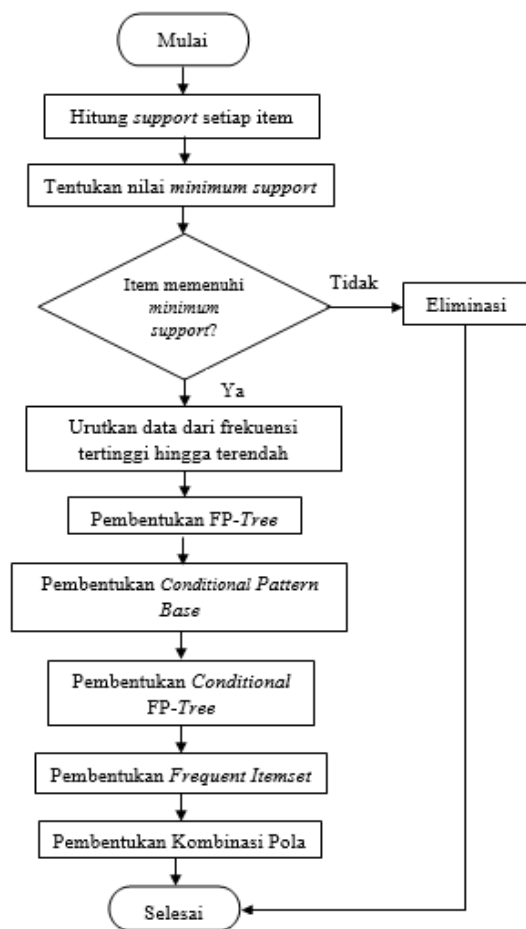
Aturan asosiasi adalah metode yang populer dalam penambangan data, tetapi penggunaannya sering menghasilkan terlalu banyak pola yang memerlukan analisis lebih lanjut untuk menemukan aturan yang menarik dan relevan.[5]

2.6. Algoritma FP-Growth

FP-growth adalah teknik data mining yang dipakai untuk menemukan pola item yang sering terjadi dalam set data transaksi. Tujuannya sama dengan association rule mining, yaitu mencari keterkaitan antara item yang terdapat dalam transaksi. Namun, FP-growth memiliki kecepatan yang lebih cepat dan membutuhkan sedikit memori dibandingkan dengan algoritma lain seperti Apriori. Oleh karena itu, FP-growth sering digunakan dalam aplikasi data mining yang memiliki dataset transaksi yang besar dan kompleks. FP-Growth ialah algoritma pengembangan dari Apriori yang dapat secara efisien menentukan frequent itemset dalam sebuah kumpulan data dengan hanya sedikit mengakses database aslinya. [6]

Berikut ini flowchart yang menjelaskan proses data mining dengan menggunakan algoritma FP-Growth.

Flowchart tersebut menjelaskan Berikut adalah langkah-langkah yang dapat dilakukan untuk menerapkan proses data mining menggunakan algoritma FP-Growth: pertama, menghitung frekuensi kemunculan dan nilai support dari setiap item dalam dataset. Kemudian, menetapkan nilai support minimum yang akan digunakan. Item yang memiliki nilai support kurang dari nilai minimum akan dihapus, sedangkan item yang memenuhi nilai minimum akan digunakan untuk tahap selanjutnya.



Gambar 1. flowchart fp-Growth

3. METODE PERANCANGAN

3.1. Sumber Pengambilan Data

Sumber data dalam tugas akhir ini diambil dari perusahaan distributor PT Dawam Lestari Jaya data yang diambil berupa data transaksi yang terjadi pada bulan Desember 2022 yang beralamat di Jl. Garuda No.16, Pekiringan, Kec. Kesambi, Kota Cirebon, Jawa Barat

3.2. Teknik Pengumpulan Data

Dalam pengumpulan data metode yang digunakan yaitu melakukan observasi secara langsung pada perusahaan melakukan wawancara dengan pihak PT Dawam Lestari Jaya untuk meneliti data-data transaksi penjualan chemicals yang terkait dengan permasalahan ini. Pada penulisan ini dilakukan untuk mengumpulkan informasi tentang literatur dan pedoman dalam penentuan hasil analisa data transaksi penjualan produk chemicals.

Sebelum dilakukannya observasi tahap pertama yang dilakukan adalah wawancara kepada pihak perusahaan meminta data transaksi yang terjadi pada perusahaan dalam satu tahun terakhir. Observasi yang dilakukan yaitu mengumpulkan berupa data transaksi perusahaan yang terjadi dalam periode satu tahun, data

perusahaan yang tersedia pada saat observasi dalam bentuk manual atau arsip yang kemudian dilakukan perapihan data dalam bentuk *excel* untuk dilakukannya pengolahan data selanjutnya.

3.3. Analisis Data

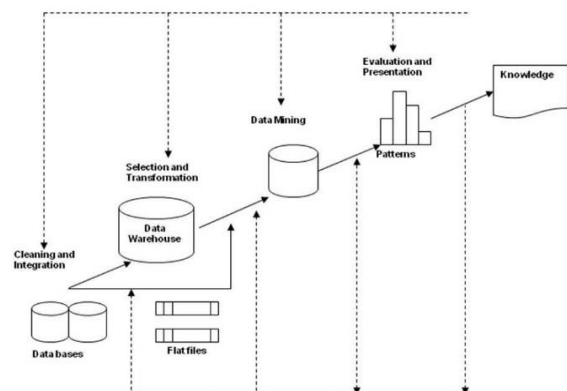
Dalam analisis ini, fokus pada penjualan produk chemicals di PT. Dawam Lestari Jaya untuk menemukan pola penjualan yang terkait. Langkah awalnya adalah menyiapkan data penjualan bulanan dari Januari hingga Desember 2022, dan kemudian data diimpor ke dalam database Microsoft Excel yang mendukung aplikasi data mining. Pemodelan menggunakan algoritma FpGrowth dan teknik data mining yang dipilih adalah asosiasi, dengan tujuan untuk menemukan aturan asosiasi dan menentukan stok produk Chemical.

3.4. Tahapan Perancangan

Terminologi "data mining" dan "knowledge discovery in database" (KDD) sering digunakan secara beralih untuk mengacu pada proses mengeksplorasi informasi yang tersembunyi dalam basis data besar. Walaupun istilah ini memiliki konsep yang berbeda, namun ada hubungan yang saling berkaitan. Data mining merupakan tahap penting dalam proses KDD.

KDD (*Knowledge Discovery in Database*) ialah proses penggalian dan analisis terhadap jumlah data besar untuk mengekstrak informasi dan pengetahuan yang berguna, yang sebelumnya tidak diketahui dan berpotensi bermanfaat, dari dalam basis data yang besar. Salah satu tahap dari serangkaian proses iteratif KDD adalah data *mining*.

Proses KDD dapat dilihat pada gambar berikut ini. [7]



Gambar 2. Proses KDD

4. HASIL DAN PEMBAHASAN

4.1. Proses implementasi

Dalam menerapkan algoritma Fp-Growth, terdapat beberapa langkah yang harus diikuti., dimulai dengan persiapan dataset yang akan diolah.

Tabel 1. Dataset Transaksi

No	Tanggal	miller	Snap Clean	Handsoap miller	Karbol Wangi	Glass cleaner	Bowl Cleaner	Odor less	Dust Cleaner	Hand Sanitizer (Gel)	Furniture Cleaner
1	01/01/2022	15	0	4	7	0	4	3	0	2	1
2	03/01/2022	8	0	1	5	0	3	0	5	0	0
3	04/01/2022	0	0	2	9	0	4	9	0	2	0
4	05/01/2022	12	5	0	4	1	0	2	2	0	0
5	06/01/2022	5	5	11	0	0	4	0	2	2	7
6	07/01/2022	3	5	0	2	4	0	7	2	0	2
7	08/01/2022	15	5	0	14	0	4	0	2	2	0
27	30/01/2022	2	4	0	2	0	0	4	4	0	4
28	31/01/2022	2	2	7	2	0	0	2	2	2	2
287	01/12/2022	2	2	0	4	2	2	2	5	0	0
288	02/12/2022	5	4	6	2	2	0	2	8	0	0
305	21/12/2022	0	5	0	0	0	1	5	2	0	6
306	26/12/2022	5	2	6	4	2	0	5	0	0	5
307	27/12/2022	0	0	5	4	5	0	8	0	6	0
308	28/12/2022	18	2	0	4	4	2	0	0	5	0
309	29/12/2022	0	9	0	2	0	2	7	0	0	2
310	30/12/2022	6	4	2	0	5	0	0	12	0	2

Tahap pertama ialah menyiapkan dataset yang akan diolah, yang terdiri dari 310 rekaman transaksi data yang disimpan dalam aplikasi Microsoft Excel dengan nama file "Data Transaksi.xls". Selanjutnya, dataset tersebut akan diuji menggunakan perangkat lunak RapidMiner untuk melihat hasil pencarian itemset frekuensi yang cocok dan relevan.



Gambar 3. Nilai settingan Confidence AssociationRules

Setting nilai confidence pada Association Rules pada rapidminer dapat mempengaruhi keakuratan dan relevansi aturan asosiasi yang ditemukan. Nilai minimal confidence menentukan tingkat kepercayaan minimum dari aturan asosiasi yang ditemukan. Nilai ini berkisar antara 0 dan 1, di mana 1 adalah tingkat kepercayaan maksimal. Misalnya, jika nilai minimal confidence ditetapkan pada 0.6, maka hanya akan ditemukan aturan asosiasi dengan tingkat kepercayaan setidaknya 0.6. Setting nilai confidence pada Association Rules pada rapidminer dapat mempengaruhi hasil akhir.

4.2. Association Rules

[Snap Clean] --> [Bowl Cleaner] (confidence: 0.601)
 [Dust Cleaner] --> [Odorless] (confidence: 0.604)
 [Hand Sanitizer (Gel)] --> [miller] (confidence: 0.607)
 [Furniture Cleaner] --> [miller] (confidence: 0.608)
 [Furniture Cleaner] --> [Dust Cleaner] (confidence: 0.608)
 [Dust Cleaner] --> [miller] (confidence: 0.610)
 [Dust Cleaner] --> [Snap Clean] (confidence: 0.610)
 [Odorless] --> [miller] (confidence: 0.610)
 [Bowl Cleaner] --> [Hand Sanitizer (Gel)] (confidence: 0.613)
 [Hand Sanitizer (Gel)] --> [Bowl Cleaner] (confidence: 0.613)
 [Snap Clean] --> [Odorless] (confidence: 0.614)
 [Glass cleaner] --> [miller] (confidence: 0.618)
 [Furniture Cleaner] --> [Odorless] (confidence: 0.620)
 [Karbol Wangi] --> [miller] (confidence: 0.629)
 [Snap Clean] --> [Dust Cleaner] (confidence: 0.633)
 [Handsoap miller] --> [miller] (confidence: 0.634)
 [Snap Clean] --> [miller] (confidence: 0.652)

Hasil dari *association rules* menunjukkan pola hubungan antara item dalam transaksi. Dari data tabel terlihat bahwa item hand sanitizer (Gel) dan chemicals Miller sering dibeli bersama, sehingga dapat dibuat aturan asosiasi bahwa jika seseorang membeli item hand sanitizer (Gel), maka kemungkinan besar mereka juga akan membeli item chemicals Miller.

No	Precondition	Conclusion	Support	Confidence	Lift
3	Hand Sanitizer (Gel)	millar	0.038	0.607	0.808
4	Furniture Cleaner	millar	0.018	0.608	0.805
5	Furniture Cleaner	Dust Cleaner	0.018	0.608	0.805
6	Dust Cleaner	millar	0.038	0.610	0.813
7	Dust Cleaner	Snap Clean	0.038	0.610	0.813
8	Odorless	millar	0.048	0.610	0.808
9	Bowl Cleaner	Hand Sanitizer (Gel)	0.048	0.613	0.807
10	Hand Sanitizer (Gel)	Bowl Cleaner	0.048	0.613	0.807
11	Snap Clean	Odorless	0.028	0.614	0.817
12	Glass cleaner	millar	0.018	0.618	0.812
13	Furniture Cleaner	Odorless	0.028	0.620	0.819
14	Karbol Wangi	millar	0.054	0.629	0.817
15	Snap Clean	Dust Cleaner	0.038	0.633	0.814
16	Handsoap miller	millar	0.044	0.634	0.811
17	Snap Clean	millar	0.048	0.652	0.808

Gambar 4. Hasil AssociationRules

Nilai confidence menunjukkan tingkat keyakinan bahwa jika item hand sanitizer (Gel) muncul, maka item chemicals Miller juga akan muncul. Hasil dari analisis ini dapat digunakan oleh perusahaan untuk menentukan strategi pemasaran, seperti menawarkan item chemicals Miller kepada pelanggan yang sering membeli item sanitizer (Gel), atau menempatkan item hand sanitizer (Gel) dan chemicals Miller berdekatan dalam toko untuk memudahkan pelanggan dalam membeli kedua item tersebut.

4.3. Hasil akhir

Menerapkan algoritma FP-Growth untuk mengidentifikasi pola pembelian produk chemicals, dapat disimpulkan dari aturan tersebut bahwa konsumen cenderung membeli produk yang saling berkaitan, misalnya jika membeli Chemicals Snap clean maka juga kemungkinan membeli chemicals bowl cleaner, dengan nilai kepercayaan (confidence) sebesar 0,6.

Hasil akhir dari aturan tersebut Analisis data menggunakan algoritma FP-Growth dapat memberikan manfaat dalam bentuk pemahaman baru mengenai pola pembelian konsumen yang sebelumnya tidak diketahui. Informasi ini dapat membantu perusahaan meningkatkan kualitas layanan kepada konsumen dengan menyediakan produk yang sesuai dengan preferensi mereka. Perusahaan harus selalu memperhatikan kebutuhan konsumen dan menyediakan produk yang diminati oleh mereka.

5. KESIMPULAN DAN SARAN

Pemelitian ini dilakukan dengan tujuan untuk mengidentifikasi pola pembelian produk chemicals yang signifikan dan relevan bagi konsumen, sehingga dapat membantu perusahaan dalam meningkatkan strategi pemasaran dan penjualan produk chemicals perusahaan. Dataset yang digunakan dalam tugas akhir ini terdiri dari 310 transaksi pembelian produk chemicals dalam periode waktu 1 tahun, yang kemudian diolah dan dianalisis dengan menggunakan algoritma Frequent Pattern-Growth.

Hasil akhir menunjukkan bahwa penerapan teknik data mining dengan algoritma Frequent Pattern-Growth mampu menghasilkan pola pembelian produk chemicals yang signifikan dan berpotensi untuk meningkatkan efektivitas strategi pemasaran dan penjualan produk chemicals. Pola pembelian tersebut mencakup informasi tentang kombinasi produk chemicals yang sering dibeli oleh konsumen, urutan pembelian produk chemicals, dan kecenderungan pembelian produk chemicals pada periode waktu tertentu. Penggunaan teknik data mining dapat membantu perusahaan untuk memahami perilaku konsumen dalam membeli produk chemicals, sehingga dapat meningkatkan kepuasan konsumen dan meningkatkan keuntungan perusahaan.

Dalam kesimpulannya, observasi pada tugas akhir ini menunjukkan bahwa penerapan teknik data mining dengan algoritma Frequent Pattern-Growth mampu menghasilkan pola pembelian produk chemicals yang signifikan dan relevan bagi konsumen.

Berdasarkan analisis dalam tugas akhir ini, terdapat beberapa rekomendasi untuk meningkatkan pengembangan sistem algoritma mining, yaitu sebagai berikut: Pihak perusahaan dapat melakukan evaluasi terhadap hasil analisis mining pada tugas akhir ini, untuk mengetahui pola pembelian customer serta memaparkan kontribusi terhadap bidang ilmu pengetahuan secara lebih luas. Selalu mempersiapkan dan merapikan data transaksi perusahaan dengan rapi dan sedetail mungkin sebelum diolah atau dianalisis dengan algoritma Frequent Pattern-Growth untuk keperluan riset selanjutnya.

DAFTAR PUSTAKA

- [1] Erasmus Simanjuntak, H. (2020). Analisa Data Mining Menggunakan Frequent Pattern Growth pada Data Transaksi Penjualan PT Mora Telematika Indonesia untuk Rekomendasi Strategi Pemasaran Produk Internet. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 4, 914–923. <https://doi.org/10.30865/mib.v4i4.2300>
- [2] Kurniawan, S., Gata, W., Wiyana, H., & Studi, P. (2018). ANALISIS ALGORITMA FP-GROWTH UNTUK REKOMENDASI PRODUK PADA DATA RETAIL PENJUALAN PRODUK KOSMETIK (STUDI KASUS: MT SHOP KELAPA GADING). In *STMIK Mercusuar Bekasi Jl. Raya Jatiwaringin* (Issue 18).
- [3] Narulita, S., Tigor Oktaga, A., & Susanti, I. (2021). *Media Aplikom Pengujian Model Prediksi Menggunakan Metode Data Mining Classification Decision Tree Untuk Penentuan Peminatan Peserta Didik*. <https://doi.org/10.33488/1.ma.2.1.305>
- [4] Nofitri, R., & Irawati, N. (2019). ANALISIS DATA HASIL KEUNTUNGAN MENGGUNAKAN SOFTWARE RAPIDMINER. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 5(2), 99–204. <https://doi.org/10.33330/jurteksi.v5i2.365> Profile. <https://linktr.ee/dawam.lj>
- [5] Wijaya. (2018). RANCANG BANGUN APLIKASI REKOMENDASI JUDUL BUKU MENGGUNAKAN ALGORITMA FP-GROWTH.
- [6] Wulan Sari, R., & Hartama, D. (2018). *Seminar Nasional Sains & Teknologi Informasi (SENSASI) Data Mining: Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi*. http://seminar-id.com/semnas_sensasi2018.html#Page|322