

ANALISIS POLA PENJUALAN SEPATU DENGAN IMPLEMENTASI ALGORITMA APRIORI DATA MINING

Dzikri Sajidan, Nana Suarna, Tati Suprpti

Teknik Informatika, STMIK IKMI Cirebon

Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

dzikrisajid023@gmail.com

ABSTRAK

Penjualan sepatu dalam industri *fashion* modern sangat vital, tetapi seringkali penjual mengalami kendala dalam pendataan dan pengelolaan stok, terutama dengan penggunaan media kertas yang rentan terhadap kehilangan data. Permasalahan tersebut mendorong penelitian untuk menerapkan algoritma apriori pada data penjualan sepatu guna mengidentifikasi jenis sepatu terlaris. Diharapkan bahwa penggunaan RapidMiner akan memberikan data yang lebih akurat dan memudahkan analisis barang karena metode Algoritma apriori digunakan untuk menentukan barang mana yang banyak dibeli dan barang mana yang kurang dibeli. Untuk mengklasifikasikan penjualan sepatu, metode data mining Algoritma Apriori digunakan. Rumus Algoritma Apriori akan digunakan untuk menghitung data yang masuk ke sistem informasi dan memberikan nilai yang lebih akurat dan valid. Kemampuan untuk mengetahui pola frekuensi asosiasi elemen adalah fase analisis yang menarik perhatian banyak peneliti saat menciptakan algoritma apriori yang efisien. dengan *support* dan *confidence* sebagai tolak ukur. Dengan demikian, penggunaan metode Algoritma apriori melalui RapidMiner diharapkan dapat meningkatkan manajemen persediaan, mempercepat proses penjualan, dan memberikan wawasan yang lebih dalam kepada penjual tentang keinginan pelanggan. Keseluruhan, penerapan algoritma ini diharapkan memberikan nilai yang lebih valid dan akurat dalam pengambilan keputusan terkait strategi bisnis dan peningkatan penjualan sepatu. Hasilnya ditemukan bahwa aturan asosiasi paling tinggi terjadi antara sepatu Sneakers Kanvas dengan nilai *support* x *confidence* sebesar 87,50%.

Kata kunci : Algoritma Apriori, Data Mining, Sepatu, Penjualan, Rapid Miner

1. PENDAHULUAN

Dalam era digital, Informatika mengubah berbagai sektor dengan pengaruhnya yang luas. Data Mining, khususnya Algoritma Apriori, menjadi fokus penelitian terkait penjualan sepatu. Sepatu bukan hanya perlindungan kaki tapi juga bagian penting dari gaya hidup dan *fashion*.

Permasalahan dijumpai dalam manajemen stok dan pemahaman data oleh pedagang sepatu, memperlambat produktivitas dan efisiensi. Data yang melimpah belum dimanfaatkan optimal untuk meningkatkan penjualan karena kurangnya sistem pendukung keputusan.

Penelitian ini bertujuan mengembangkan sistem klasifikasi penjualan sepatu berbasis Data Mining menggunakan Algoritma Apriori melalui RapidMiner, dengan harapan model yang dihasilkan dapat mengoptimalkan data, memproses penjualan secara efisien, dan membantu pendataan stok barang. Metode Algoritma Apriori diharapkan dapat mengatasi masalah di toko sepatu, sementara Penggunaan RapidMiner diharapkan dapat memberikan data keputusan yang lebih akurat dan mempermudah analisis barang. Dengan penerapan data mining menggunakan metode Algoritma Apriori dalam mengklasifikasi Penjualan Sepatu, dimana nantinya data yang masuk ke sistem informasi dihitung dengan rumus Algoritma Apriori yang Hasilnya akan disajikan secara rinci, mampu memberikan nilai yang valid dan lebih akurat. Salah satu fase analisis yang menarik perhatian banyak peneliti dalam menciptakan

algoritma apriori yang efisien adalah kemampuan untuk mengetahui pola frekuensi asosiasi suatu elemen. dengan dua tolak ukur yaitu *Support* dan *Confidence*. Penggunaan Algoritma Apriori melalui RapidMiner juga membantu mengidentifikasi barang yang diminati dan kurang diminati, memberikan wawasan berharga bagi peneliti dan praktisi di bidang Informatika.

2. TINJAUAN PUSTAKA

Setelah peneliti melakukan telaah terhadap beberapa penelitian, ada beberapa yang memiliki keterkaitan dengan penelitian yang peneliti lakukan. Penelitian yang berhasil peneliti temukan adalah Berdasarkan hasil Literature review dan perbandingan algoritma atau skema yang diterapkan pada beberapa paper dibawah ini:

Penelitian sebelumnya membahas tentang Meskipun perusahaan telah menerapkan teknologi informasi, potensi data transaksi penjualan seringkali belum dimaksimalkan, hanya digunakan dalam laporan rutin. Kebutuhan akan informasi akurat menuntut penambangan data atau KDD dengan teknik statistik, kecerdasan buatan, dan pembelajaran mesin. PT Global Vision, pemasok sepatu, masih menggunakan metode manual dalam strategi penjualan, kurangnya pedoman khusus dan pengelolaan data yang baik dapat menghambat kemajuan perusahaan. Untuk meningkatkan efisiensi strategi penjualan, perusahaan perlu mempertimbangkan penerapan teknologi dan analisis

data yang lebih canggih. Dengan penggunaan yang tepat, data yang terus bertambah dapat diolah menjadi wawasan bernilai, menghindari akumulasi data yang tak bermanfaat dan merugikan. [1]

Adapun penelitian tentang CV Princes Diary, perusahaan aksesoris dengan beragam produk, ingin meningkatkan analisis data penjualan untuk strategi penjualan. Melalui penelitian ini, dikembangkan aplikasi Data Mining dengan algoritma Apriori menggunakan RapidMiner, menghasilkan enam aturan strategi penjualan. Contohnya, hubungan antara pembelian kacamata dan dompet dengan tingkat *support* 52,5% dan *confidence* 55,9%. Temuan tertinggi adalah pembelian make up diikuti pembelian sisir, *support* 8,825% dan *confidence* 0,825%. [2]

Penjualan sepatu menjadi tren yang merangkul berbagai kalangan, namun dalam toko seringkali muncul kendala dalam mengidentifikasi sepatu favorit pelanggan. Penelitian ini menggunakan Metode *K-Means Clustering* untuk menganalisis dan mengelompokkan tipe sepatu yang terjual di toko Ritelindo. Dalam analisis ini, terbentuk tiga kelompok cluster: Terlaris, Sedang, dan Terendah. Hasilnya menunjukkan 4 item masuk ke kelompok Terlaris, 12 item ke kelompok Sedang, dan 9 item ke kelompok Terendah. Dengan pendekatan ini, pemilik toko dapat lebih jelas mengidentifikasi sepatu yang diminati pelanggan dan menyesuaikan strategi penjualan untuk meningkatkan efektivitasnya. [3]

Penelitian selanjutnya membahas tentang PT. Sepatu Bata menggunakan Prediksi Penjualan dengan Regresi Linier Sederhana. Analisis ini menentukan hubungan linier antara variabel bebas dan terikat, memproyeksikan nilai terikat saat bebas mengalami perubahan. Penelitian ini bertujuan memprediksi penjualan sepatu populer dengan data transaksi yang ada. Sistem data mining yang dikembangkan menggunakan PHP dan MySQL untuk implementasi. Melalui metode ini, PT. Sepatu Bata berupaya memperoleh wawasan yang lebih mendalam terkait penjualan sepatu dan mempersiapkan strategi yang lebih efektif berdasarkan pemahaman yang diperoleh dari analisis data historis. [4]

Penelitian ini menyoroti peran data mining dalam proyeksi penjualan sepatu merek Adidas melalui algoritma regresi linier. Sepatu bukan hanya pelindung kaki, tapi juga ekspresi gaya. Metode kualitatif digunakan dengan kajian literatur dari jurnal-jurnal terkait, referensi diperoleh dari Google Scholar, Neliti, dan sumber online lainnya. Hasilnya menunjukkan bahwa data mining berpengaruh signifikan dalam memprediksi penjualan sepatu Adidas. Algoritma regresi linier sederhana efektif dalam analisis prediksi, menemukan persamaan regresi $Y = 122.666,5 + 12.008,6X$, dengan tingkat akurasi sebesar 133.978,6%. Temuan ini menegaskan pentingnya data mining dalam proyeksi penjualan sepatu dan keberhasilan algoritma regresi linier sebagai metode analisis yang kuat [5]

Penelitian ini fokus pada PT Global Vision, pemasok sepatu yang ingin meningkatkan omset melalui strategi bisnis yang lebih terstruktur. Mereka menggunakan algoritma FP-Growth pada data transaksi penjualan sepatu untuk mengidentifikasi pola pembelian bersama. Metode *Association Rule* dan *FP-Tree* digunakan untuk menganalisis kombinasi produk, memberikan wawasan tentang merk sepatu yang sering dibeli bersamaan oleh konsumen. Hasilnya, merk sepatu V-Kids memiliki keterkaitan tinggi dengan merk Vitamori (0,9), sementara Canvas memiliki keterkaitan sebesar 0,5 dengan keduanya. Penelitian merekomendasikan penggunaan sumber data yang lebih besar untuk aturan asosiasi yang lebih kaya dan eksplorasi algoritma lain seperti eclat. Penerapan FP-Growth memberikan potensi bagi PT Global Vision dalam pengambilan keputusan dan strategi stok yang lebih optimal, sementara penelitian lanjutan bisa mempertimbangkan aplikasi datamining lainnya untuk perbandingan hasil dan keakuratan perhitungan [6]

Penelitian ini menggunakan data mining dengan algoritma *Clustering K-Means* dan aplikasi Rapidminer untuk menganalisis minat konsumen terhadap produk ecoprint berdasarkan beberapa variabel. Hasilnya, data terbagi menjadi 3 cluster: sangat diminati, diminati, dan kurang diminati. Sepatu sport dan sepatu flat wanita diminati secara signifikan, sementara sepatu flat pria dan sepatu boots memiliki tingkat minat yang rendah. Preferensi juga dipengaruhi oleh jenis kelamin, usia, dan daerah. Wanita, usia 25-50 tahun, terutama dari Pulau Jawa dan sekitarnya, menunjukkan minat yang tinggi. Informasi ini menjadi panduan bagi produsen ecoprint untuk memahami preferensi konsumen dan merancang strategi pemasaran yang lebih efektif. [7]

Penelitian pada Toko Sepatu Kaki-kaki menggunakan metode *K-Means Clustering* untuk mengidentifikasi tren produk yang diminati oleh konsumen. Hasilnya menunjukkan 2 produk sangat diminati, 7 produk cukup populer, dan 1 produk kurang diminati. Analisis ini memberikan dasar yang kuat berdasarkan data transaksi penjualan untuk pengambilan keputusan yang lebih informasional. Dengan wawasan ini, toko dapat lebih tepat dalam menyesuaikan stok produk dengan tren pasar dan preferensi pelanggan, mengurangi risiko yang terkait dengan ketidakpastian tren dan kebutuhan konsumen. [8]

Penelitian berikutnya membahas PT. Pratama Abadi Industri, perusahaan sepatu, menghadapi kendala persediaan bahan baku yang mempengaruhi produksi sepatu paling diminati. Dengan menggunakan metode Algoritma Apriori dan aplikasi Tanagra 1.4.50, penelitian ini menganalisis pola pemesanan sepatu Nike. Tujuannya adalah mendapatkan aturan asosiasi antar item dengan nilai minimum *support* 30% dan minimum *confidence* 70%. Hasil analisis dari Mei hingga Oktober 2020 menunjukkan pola pemesanan sepatu Nike yang paling

diminati, seperti Tailwind (100%), Revolution 5 (100%), Air Max Plus (100%), dan Classic Cortez Leather (100%). Informasi ini dapat digunakan untuk mengoptimalkan persediaan bahan baku, memungkinkan PT. Pratama Abadi Industri lebih efisien dalam memenuhi permintaan pelanggan dan menjaga kelancaran proses produksi. [9]

Adapun penelitian dengan judul "Implementasi Data Mining Algoritma Apriori untuk Meningkatkan Penjualan" membahas Lakoe Dessert Pondok Kacang, dengan 45 jenis produk, melakukan analisis data penjualan menggunakan algoritma Apriori dan aplikasi Rapidminer. Dengan nilai support 2,4% dan confidence 50%, penelitian ini mengidentifikasi Pudding Strawberry dan Pudding Vanilla sebagai produk paling banyak dibeli oleh konsumen pada bulan Maret 2020. Analisis ini memungkinkan perusahaan mengembangkan strategi pemasaran yang lebih baik, memasarkan produk lain berdasarkan keuntungan dari penjualan produk terlaris, dan mengantisipasi potensi kehabisan stok atau bahan. [10]

3. METODE PENELITIAN

Dalam penelitian ini, dipilih metode *Association Rules*, suatu algoritma yang bertujuan untuk mengidentifikasi pola yang sering terjadi di dalam sejumlah besar transaksi. Setiap transaksi terdiri dari sejumlah item, sehingga pendekatan ini mendukung pengembangan sistem rekomendasi dengan menemukan pola asosiasi antar item dalam transaksi yang tercatat. Untuk mendapatkan nilai *support*, dapat menggunakan persamaan berikut:

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

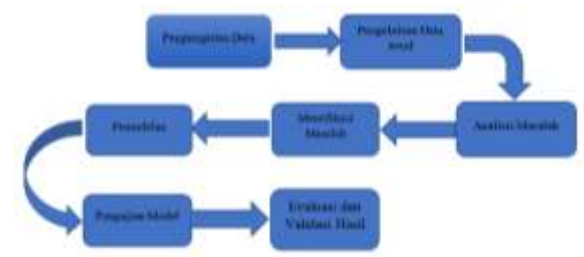
Selanjutnya, langkah berikutnya adalah menentukan rule dengan mencari *confidence* dari frequent itemset tersebut, menggunakan persamaan berikut:

Confidence

$$P/(B \setminus A) = \frac{\text{Jumlah transaksi A dan B}}{\text{Jumlah transaksi mengandung A}} \quad (2)$$

Support atau nilai penunjang adalah persentase kombinasi item tersebut dalam database, sedangkan *confidence* atau nilai kepastian adalah kuatnya hubungan antar item dalam aturan asosiasi. Algoritma Apriori dapat membantu dalam transaksi penjualan bisnis maupun bidang ritel.

Penelitian ini juga menggunakan metode algoritma FG-Growth. Algoritma secara umum diartikan sebagai serangkaian langkah atau instruksi yang jelas untuk menyelesaikan masalah. Dalam pemrograman, algoritma adalah kumpulan instruksi sistematis yang digunakan untuk menanggulangi hambatan logika atau matematika dengan dukungan komputer. Tahapan yang akan digunakan dalam penelitian ini dapat dilihat pada Gambar 1 untuk melakukan Klasifikasi Prediksi terhadap data hasil Penjualan Sepatu. Berikut ini beberapa langkah dalam tahapan penelitian yang akan dilakukan.



Gambar 1. Metode Penelitian

3.1. Sumber Data

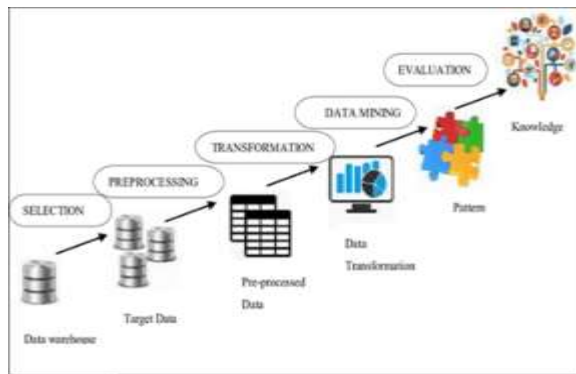
Data primer yang diperoleh merupakan sumber informasi langsung yang disediakan oleh partisipan atau objek penelitian kepada peneliti. Dalam konteks penelitian ini, data primer diperoleh melalui pelaksanaan survei dan observasi langsung ke toko Aren Sepatu. Berikut contoh data seperti pada Tabel 1 berikut:

Table 1. Data Penjualan

NO	Pembeli	Tanggal Transaksi	Ukuran	Brand Sepatu	Jenis Sepatu
1	Ahmad	01/01/2022	42	Ventela	Sneakers
2	Ardi	01/01/2022	40	Adidas Stan Smith	Sneakers
3	Sri	01/01/2022	38	New Balance	Sport
4	Thiara	01/01/2022	37	Dr Martens	Pantofel
5	Didin	01/01/2022	40	Mizuno	Sport
.....					
.....					
250	Iim	09/06/2022	38	Adidas Stan Smith	Sneakers

3.2. Teknik Analisis Data

Adapun untuk menganalisis data dalam penerapan data mining ini menggunakan proses tahapan *Knowledge Discovery in Databases (KDD)* yang terdiri dari, *Data Cleaning*, *Data transformation*, *Data mining*, *Pattern evolution*, dan *knowledge*. tampilan KDD bisa dilihat pada gambar 2



Gambar 2. Metode KDD

1. Selection (seleksi)

Untuk membaca dataset dalam bentuk file excel, menggunakan operator Read Excel. Data yang digunakan pada penelitian ini ialah sebanyak 250 record dataset.

2. Preprocessing (pemilihan data)

Proses *cleaning* atau pembersihan data yang missing atau memiliki nilai yang tidak konsisten pada langkah *preprocessing*. Sebelum melakukan proses ini, dilakukan analisa terlebih dahulu apakah dataset memiliki nilai missing atau tidak serta memiliki data yang konsisten atau tidak.

3. Data transformation

adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Pada langkah *transformation* dilakukan untuk data bertipe polynominal menjadi numerik agar data dapat diolah berdasarkan jarak. Untuk mengubah data bertipe polynominal menjadi numerik menggunakan operator Nominal to Numerical.

4. Data mining

Data mining merupakan tahapan mengeksplorasi dan menganalisa data dalam jumlah besar yang tujuannya agar dapat mengidentifikasi suatu pola atau informasi yang menarik dari data yang tersimpan dalam jumlah yang besar dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma yang tepat sangat bergantung pada tujuan dan proses *knowledge discovery in databases* (KDD) secara keseluruhan. Pada tahapan ini, esensi dari proses *knowledge discovery in databases* (KDD) dijalankan dengan tujuan menganalisis data yang sudah melewati tahap pembersihan.

5. Evaluation (evaluasi)

Hasil dari data mining, yakni alur informasi yang muncul dari proses, perlu disajikan dalam format yang dapat dipahami dengan mudah oleh para pemangku kepentingan. Proses ini merupakan bagian dari *Knowledge Discovery in Databases* (KDD) yang mencakup penelitian terhadap apakah alur atau informasi yang ditemukan sesuai dengan fakta atau hipotesis yang telah ada sebelumnya.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Hasil penelitian yang dituangkan didalam pembahasan ini yaitu akan memaparkan proses bagaimana menganalisis dataset penjualan sepatu dengan menggunakan Algoritma apriori. penelitian ini dilakukan melalui tahap pengujian machine learning yaitu Software RapidMiner Studio.

4.2. Lokasi penelitian

Penelitian ini dilakukan di TOKO AREN SEPATU yang terletak di Jl. Sukajadi No.46, Pasteur, Kec. Sukajadi, Kota Bandung, Jawa Barat 40162.

4.3. Data selection

Data yang digunakan dalam penelitian ini adalah data Penjualan Sepatu. Kemudian data tersebut diseleksi dan akan digunakan untuk diolah menggunakan aplikasi/tools Rapidminer. Tahapan seleksi dilakukan di Microsoft Excel dengan atribut yang digunakan diantaranya No, Pembeli, Tanggal Transaksi, Jenis Sepatu, No ukuran dan Qty Jenis. Hasil seleksi data Seperti pada Tabel 2 berikut :

Table 2. Atribut default sepatu

No	Atribut	Type Data	Keterangan
1.	Pembeli	Integer	Nama Pembeli
2.	Tanggal Transaksi	Polynominal	Tanggal Transaksi
3.	Jenis Sepatu	Integer	Jenis yang ditawarkan
4.	No Ukuran	Polynominal	Ukuran Sepatu
5.	Qyt Jenis	Polynominal	Jumlah pembelian

4.4. Data preprocessing

Pada tahap ini dilakukan preprocessing data yaitu proses cleaning atau pembersihan data. Proses penghapusan data atribut yang tidak diperlukan seperti No, No ukuran dan Tanggal Transaksi. Sedangkan atribut yang akan digunakan adalah Pembeli, Jenis dan Qty Jenis (Quantity Jenis). walaupun pada atribut masih ada lagi beberapa yang dapat dijadikan atribut baru tetapi dapat diabaikan karena banyak memiliki nilai missing. Sehingga dari 5 atribut menjadi seperti tampak pada tabel 3 berikut:

Table 3. Atribut default dari file sepatu

No.	Atribut	Type Data	Keterangan
1.	Pembeli	Integer	Nama Barang
3.	Jenis Sepatu	Integer	Harga asli
5.	Qyt Jenis	Polynominal	Jumlah pembelian

4.5. Data Transformation

Pada fase Data Transformation, dilakukan langkah-langkah untuk mengubah format data agar sesuai dengan kebutuhan proses pemodelan dalam data mining. Tujuannya adalah untuk mempermudah koordinasi data yang diproses oleh algoritma dan alat yang digunakan dalam penelitian, khususnya RapidMiner. Pada tahap transformasi data penjualan,

terdapat tiga atribut utama, yaitu Pembeli (dirubah menjadi No id), Jenis, dan Qty Jenis. Atribut Jenis sendiri terdiri dari 10 sub-atribut, yakni Boots, Heels, Kanvas, Monk Strap, Oxpord, Pantofel, slip On, Sneakers, Sport dan Wedges. Proses transformasi pada tahap ini fokus pada penyesuaian struktur data agar kompatibel dengan kebutuhan algoritma dan tools yang digunakan, sehingga memungkinkan analisis data penjualan dengan lebih efisien. Data diubah kedalam bentuk tabular, seperti dalam Gambar 3 berikut:

No ID	Boots	Heels	Kanvas	Monk Strap	Oxpord	Pantofel	Slip On	Sneakers	Sport	Wedges
110	0	0	0	0	0	0	0	1	0	0
111	0	0	0	0	0	0	0	1	0	0
112	0	0	0	0	0	0	0	1	0	0
113	0	0	0	0	0	1	0	0	0	0
114	0	0	0	0	0	0	0	0	1	0
115	0	0	0	0	0	0	0	1	0	0
116	0	0	0	0	0	0	0	0	1	0
117	0	0	1	0	0	0	0	0	0	0
118	0	0	1	0	0	0	0	0	0	0
119	0	0	0	0	0	0	1	0	0	0
120	0	0	0	0	0	0	1	0	0	0
121	1	0	0	0	0	0	0	0	0	0
122	0	1	0	0	0	0	0	0	0	0
123	0	0	0	0	0	0	1	0	0	0
124	0	0	0	0	0	0	1	0	0	0
125	0	0	0	0	0	0	0	1	0	0
126	0	0	0	0	0	0	0	0	1	0
127	0	0	0	0	0	1	0	0	0	0
128	0	0	0	0	0	0	0	0	1	0
129	0	0	0	0	1	0	0	0	0	0
130	0	1	0	0	0	0	0	0	0	0

Gambar 3. Data transformation

4.6. Data Mining

Tahapan Data Mining yang diterapkan dalam penelitian ini adalah Algoritma Apriori. Algoritma Apriori dimanfaatkan untuk mengidentifikasi keterkaitan an di antara data atau bagaimana kelompok data tertentu memengaruhi keberadaan data lainnya. Pendekatan ini berguna untuk mengungkap pola-pola khusus dalam dataset yang besar. Terdapat dua tahap dalam Algoritma Apriori:

- Pencarian Frequent Itemset:** Tahap ini melibatkan pencarian kombinasi yang paling sering terjadi dari suatu itemset. Frequent itemset adalah kelompok item yang sering muncul bersama dalam transaksi.
- Penyusunan Rules:** Setelah mendapatkan frequent itemset, langkah berikutnya adalah menyusun aturan (rules). Ini dilakukan dengan mendefinisikan Condition dan Result (conditional association rule).

Dalam menghasilkan suatu *association rule*, diterapkan suatu *interestingness measure* (ukuran kepentingan) yang diperoleh melalui proses pengolahan data dengan perhitungan khusus. Terdapat tiga ukuran umum yang digunakan:

- Support**
Ukuran yang mengindikasikan sejauh mana suatu item atau set item mendominasi transaksi secara keseluruhan. Support menentukan apakah suatu item atau itemset layak untuk dicari confidence-nya.
- Confidence**
Ukuran yang mencerminkan keterkaitan antara dua item dalam konteks kondisional. Contohnya, sejauh mana item B sering dibeli ketika seseorang membeli item A.

c. Improvement

Ukuran yang mencerminkan probabilitas besar bahwa dua item akan dibeli bersamaan.

Ketiga ukuran ini membantu untuk menentukan *interesting association rule*, yang kemudian disamakan dengan *threshold* (batasan) yang ditetapkan. Batasan tersebut umumnya mencakup minimal *support*, minimal *confidence*, dan minimal *improvement*. Sebuah *association rule* dengan *confidence* sama atau lebih besar dari minimal *confidence* Y dapat dianggap sebagai valid *association rule*. Berikut pengaplikasian metode *association rule* dalam aplikasi/tools Rapidminer Diantaranya sebagai berikut :

4.7. Tampilan halaman utama Rapidminer

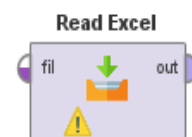
Pada gambar dibawah adalah tampilan halaman utama Rapidminer ketika dibuka



Gambar 4. Halaman utama

RapidMiner adalah salah satu software untuk pengolahan data mining. RapidMiner text mining menjalankan tugas-tugas terkait analisis teks, mengeksplorasi pola-pola dalam dataset yang besar, dan mengintegrasikannya dengan metode statistika, kecerdasan buatan, dan basis data. Fokus analisis teks ini adalah untuk memperoleh informasi berkualitas tinggi dari teks yang diproses. RapidMiner menyediakan serangkaian prosedur untuk data mining dan machine learning, termasuk ETL (*extraction, transformation, loading*), preprocessing data, visualisasi, pemodelan, dan evaluasi. Tahapan data mining diatur menggunakan operator-operator yang dapat disusun bersarang, dijelaskan melalui XML, dan dibentuk dengan antarmuka pengguna grafis (GUI). Implementasinya menggunakan bahasa pemrograman Java.

4.8. Proses input data dengan Format excel



Gambar 5. Read excel

Pada gambar diatas menjelaskan proses penginputan data. Data yang diinput bisa berbagai

macam format seperti CSV, Excel dan masih banyak lainnya. Akan tetapi file yang akan diuji di aplikasi Rapidminer ini menggunakan data dengan format Excel yang sudah disiapkan. Dalam RapidMiner, terdapat fitur Read Excel yang berperan sebagai alat pembaca file Excel. Read Excel merupakan operator yang sangat dasar dan umumnya digunakan sebagai langkah awal sebelum memulai proses lebih lanjut. Dalam konteks penelitian ini, digunakan file Excel yang memiliki format .xls.

Parameter pada operator Read Excel diterapkan dengan menggunakan pengaturan bawaan (default). Data yang diperoleh dari operasi pembacaan oleh operator Read Excel mencakup informasi sebagaimana tergambar dalam Tabel 3 di bawah ini.

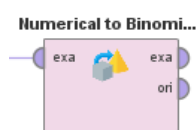
Table 4. Statik dataset

No	Uraian	Keterangan
1.	Record	250
2.	Special Attribute	0
3.	Reguler Attribute	11
	Attribute :	
1.	Jenis Sepatu	Integer

Gambar 6. Hasil Statistik

Dataset tersebut memiliki 0 special attribute, 11 regular attribute dan 250 record.

4.9. Numerical to Binominal



Gambar 7. Numerical to Binominal

Pada gambar di atas, Operator "Numerik ke Binominal" mengubah jenis atribut numerik menjadi jenis binominal (atau biner) dengan memetakan nilai-nilai numerik ke nilai binominal yang sesuai. Atribut binominal hanya memiliki dua nilai mungkin: 'benar' atau 'salah'. Operator ini melakukan pemetaan nilai-nilai atribut berdasarkan batasan minimal (min) dan maksimal (max) yang ditentukan. Secara khusus, batasan defaultnya ditetapkan ke 0,0. Artinya, jika nilai atribut berada di luar rentang antara 0 dan 0, atribut tersebut akan diklasifikasikan sebagai 'benar'. Namun, jika nilainya berada dalam rentang antara nilai minimal dan maksimal yang ditentukan (dalam kasus default, 0 dan 0), maka atribut tersebut akan

diklasifikasikan sebagai 'salah'. Sebagai contoh, jika nilai atribut adalah 2, itu akan diklasifikasikan sebagai 'salah' karena nilainya berada di dalam rentang antara nilai minimal dan maksimal yang ditentukan (0 dan 0). Namun, jika nilainya adalah 0 atau di luar rentang 0 dan 0, maka akan diklasifikasikan sebagai 'benar'.

Operator ini berguna untuk mengubah atribut numerik menjadi representasi biner yang dapat digunakan dalam berbagai analisis atau model pemrosesan data yang memerlukan jenis atribut biner. Parameter yang digunakan seperti pada Tabel 4. Hasil dari penggunaan operator ini diperoleh informasi :

Table 5. Hasil Penggunaan Operator

No.	Nama Atribut	Jenis Data	Missing
1	No Id	Polynominal	1
2	Sneakers	Binominal	0
3	Kanvas	Binominal	0
4	Heels	Binominal	0
5	Monk Strap	Binominal	0
6	Pantofel	Binominal	0
7	Sport	Binominal	0
8	Slip On	Binominal	0
9	Wedges	Binominal	0
10	Oxford	Binominal	0

4.10. Fp-Growth



Gambar 8. Fp-Growth

FPGrowth (Frequent Pattern Growth) merupakan satu dari algoritma yang dipergunakan agar dapat mengidentifikasi pola-pola yang kerap muncul dalam data transaksional, seperti dataset penjualan sepatu. Dalam RapidMiner, pengaplikasian algoritma FP-Growth mencakup deteksi asosiasi antara sepatu sepatu yang kerap dibeli bersama-sama didalam data penjualan sepatu. Algoritma inipun memungkinkan pemakai supaya dapat menemukan pola transaksi penjualan yang sering muncul dalam dataset, memberikan pengetahuan tentang kecenderungan atau hubungan yang signifikan antara item-item tertentu. Parameter yang digunakan seperti dalam tabel 5 dibawah. Hasil dari penggunaan operator ini diperoleh informasi :

Table 6. Parameter Fp Growth

Size.	Support	Item 1	Item 2
1	0.047	Slip on	
2	0.075	Pantofel	Monk strap
2	0.082	Sneakers	Kanvas

4.11. Create Association Rule

Pada gambar 9 operasi "Create Association Rules" dipergunakan untuk membuat aturan asosiasi dari data. Aturan asosiasi menggambarkan hubungan atau pola yang muncul bersamaan dalam dataset, seperti sepatu yang kerap dibeli bersama dalam

transaksi penjualan atau item yang kerap muncul dalam urutan tertentu. Parameter yang digunakan seperti dalam Gambar 10 berikut. Hasil dari penggunaan operator ini diperoleh informasi :



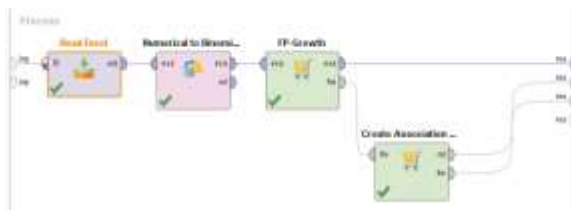
Gambar 9. Create Association rule

No	Premis	Conclusion	Support	Confidence	LaPlace	Gain	p-a	Lift	Conviction
1.	Sneakers	Kanvas	0.082	0.875	0.969	-0.106	-0.038	1.839	4.235

Gambar 10. Parameter Create Assosiation Rule

4.12. Apply Model

Apply model merupakan tahapan akhir yang mana pada tahap ini akan menerapkan model data mining Asosiasi berdasarkan proses-proses sebelumnya mulai dari Read excel, Numerical to binominal, FP-Growth dan Create Association Rule. Setelah semua proses dilakukan, maka akan didapatkan hasil model Asosiasi dari data testing sebesar 20% seperti pada gambar 11 dibawah ini.



Gambar 11. Apply model

Pada gambar di atas di bagian operator Read Excel data akan dimasukan seperti gambar 12 di bawah ini.

	A	B	C	D	E	F	G	H	I	J
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
12	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
13	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Gambar 12. Apply model

Hasil Algoritma Apriori/ Association Rules yang sudah di hitung melalui rapid miner.

[Sneakers] --> [Slip On, Boots] (confidence: 0.588)
[Wedges] --> [Sport] (confidence: 0.600)
[Monk Strap] --> [Sport] (confidence: 0.600)
[Slip On] --> [Sport, Heels] (confidence: 0.667)
[Sport] --> [Heels, Boots] (confidence: 0.667)
[Sport] --> [Pantofel, Boots] (confidence: 0.706)
[Pantofel] --> [Sport, Monk Strap] (confidence: 0.714)
[Sneakers] --> [Sport, Kanvas] (confidence: 0.750)
[Heels] --> [Sneakers, Sport] (confidence: 0.750)
[Pantofel] --> [Sneakers, Kanvas] (confidence: 0.800)
[Boots] --> [Sneakers, Sport] (confidence: 0.800)
[Sport] --> [Sneakers] (confidence: 0.833)
[Boots] --> [Sneakers, Pantofel] (confidence: 0.857)
[Pantofel] --> [Monk Strap] (confidence: 0.857)
[Sneakers] --> [Kanvas] (confidence: 0.875)

Hasil operasi Algoritma Apriori/Association Rules berupa peraturan asosiasi yang melibatkan kombinasi item atau atribut yang seringkali muncul bersama-sama, disertai tingkat kepercayaan (*Confidence*) dan dukungan (*Support*) untuk setiap peraturan tersebut.

Contoh penerapan Association Rules dapat terlihat dalam analisis transaksi penjualan, misalnya peraturan yang menunjukkan bahwa "Jika pelanggan membeli sepatu Sport, maka kemungkinan besar pelanggan tersebut juga akan membeli sepatu hak tinggi dan sepatu boots."

5. KESIMPULAN DAN SARAN

Kesimpulan dari analisis pola penjualan sepatu dengan menerapkan algoritma Apriori dalam data mining dapat dirangkum sebagai berikut: Penemuan Pola Keterkaitan Penjualan Melalui penerapan algoritma Apriori pada 250 transaksi penjualan sepatu, ditemukan pola-pola terkait penjualan dengan menetapkan nilai minimum *support* sebesar 0.05 dan tingkat kepercayaan (*confidence*) sebesar 0.09. Pendekatan ini memungkinkan identifikasi kombinasi yang memiliki signifikansi didalam konteks penjualan sepatu. Penilaian Kuatnya Aturan Asosiasi Aturan asosiasi ini memberikan wawasan kepada pemilik toko untuk mengatur penataan barang atau merancang kampanye pemasaran. Berdasarkan nilai *support x confidence*, ditemukan bahwa aturan asosiasi paling tinggi terjadi antara sepatu Sneakers Kanvas dengan nilai *support x confidence* sebesar 87,50%. Selain itu aturan asosiasi tertinggi berikutnya antara Pantofel dan

Monk Strap memiliki nilai *support* x *confidence* sebesar 85,70%. Informasi ini dapat digunakan untuk menentukan jenis sepatu apa yang paling laris di pasaran. Informasi yang didapat dari hasil Data Mining sebagai berikut. Ketika membeli Sneakers maka akan membeli juga Kanvas dengan nilai :

Support : 0.082

Confidence : 0.875

Catatan:

a. Keyakinan (*Confidence*);

Confidence dari suatu aturan didefinisikan sebagai $\text{conf}(X \text{ menyiratkan } Y) = \frac{\text{supp}(X \cup Y)}{\text{supp}(X)}$. Keyakinan ini memiliki rentang antara 0 hingga 1. Keyakinan merupakan estimasi dari Probabilitas (Pr) dari Y, jika diberikan X. Dukungan (*support*) dari suatu itemset X didefinisikan sebagai proporsi transaksi dalam kumpulan data yang mengandung itemset tersebut.

b. Kenaikan (*Lift*):

Kenaikan dari suatu aturan ditetapkan sebagai $\text{lift}(X \text{ menyiratkan } Y) = \frac{\text{supp}(X \cup Y)}{(\text{supp}(Y) \times \text{supp}(X))}$ atau rasio dukungan yang diamati dengan yang diharapkan jika X dan Y independen. Juga dapat ditetapkan sebagai $\text{lift}(X \text{ menyiratkan } Y) = \frac{\text{conf}(X \text{ menyiratkan } Y)}{\text{supp}(Y)}$. *Lift* mengukur sejauh mana X dan Y bersifat independen. Rentangnya berkisar antara 0 hingga tak terhingga positif. Nilai mendekati 1 menunjukkan bahwa X dan Y independen dan aturan tersebut kurang menarik.

c. Keyakinan (*Conviction*):

Conviction sensitif terhadap arah aturan, yaitu $\text{conv}(X \text{ menyiratkan } Y)$ tidak sama dengan $\text{conv}(Y \text{ menyiratkan } X)$. Keyakinan ini agak terinspirasi oleh definisi logis dari implikasi dan upaya untuk mengukur tingkat implikasi suatu aturan. Keyakinan didefinisikan sebagai $\text{conv}(X \text{ menyiratkan } Y) = \frac{(1 - \text{supp}(Y))}{(1 - \text{conf}(X \text{ menyiratkan } Y))}$.

Saran dari penulis dalam penelitian ini adalah perlunya membandingkan dengan algoritma lain untuk menganalisa dataset penjualan sehingga diharapkan: Dalam asosiasi penjualan yang dilakukan oleh penulis dapat memiliki nilai akurasi yang lebih baik dan akurat. Dengan Analisa dataset penjualan dengan menggunakan algoritma apriori ini semoga selanjutnya dapat membandingkan dengan algoritma yang lain seperti : K-means, C-Means, C. 45, dan lain-lain

DAFTAR PUSTAKA

- [1] S. Syahriani, "Penerapan Data Mining Untuk Menentukan Pola Penjualan Sepatu Menggunakan Metode Algoritma Apriori," *BINA Insa. ICT J.*, vol. 9, no. 1, 2022, doi: 10.51211/biict.v9i1.1758.
- [2] S. Qomariah, B. Basrie, and S. F. Pa'a, "IMPLEMENTASI ALGORITMA APRIORI PADA DATA PENJUALAN PRODUK ASESORIS CV PRINCES DIARY SAMARINDA," *Just TI (Jurnal Sains Terap. Teknol. Informasi)*, vol. 12, no. 2, 2020, doi: 10.46964/justti.v12i2.321.
- [3] F. A. Bramasta and R. Halilintar, "Penerapan Data Mining Untuk Menentukan Strategi Penjualan Toko Sepatu," *Pros. SEMNAS INOTEK ...*, 2021.
- [4] D. A. Pohan, M. H. Dar, and Irmayanti, "Penerapan Data Mining untuk Prediksi Penjualan Produk Sepatu Terlaris Menggunakan Metode Regresi Linier Sederhana," *J. Nas. Inform. dan Teknol. Jar.*, vol. 2, 2022.
- [5] S. Supardi *et al.*, "Peran Data Mining dalam Memprediksi Tingkat Penjualan Sepatu Adidas Menggunakan Metode Algoritma Regresi Linear Sederhana," *J. Ekon. MANJEMEN Sist. Inf.*, vol. 4, 2023.
- [6] S. Syahriani, "Penerapan Data Mining Dalam Menentukan Pola Penjualan Sepatu Menggunakan Algoritma FP-Growth," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 6, 2022, doi: 10.30865/jurikom.v9i6.5148.
- [7] L. Yusuf and S. Sudirman, "Klasifikasi Kebutuhan Produk Ecoprint Menggunakan Metode Clustering K-Means," *J. Larik Ldng. Artik. Ilmu Komput.*, vol. 2, no. 2, 2023, doi: 10.31294/larik.v2i2.1797.
- [8] W. W. Kristianto, "Penerapan Data Mining Pada Penjualan Produk Menggunakan Metode K-Means Clustering (Studi Kasus Toko Sepatu Kakikaki)," *J. Pendidik. Teknol. Inf.*, vol. 5, no. 2, 2022, doi: 10.37792/jukanti.v5i2.547.
- [9] R. Wati, I. Imron, and A. Riyandi, "ANALYSIS OF SHOE ORDERING PATTERN AT PT. PRATAMA ABADI INDUSTRI USING THE APRIORI ALGORITHM," *JITK (Jurnal Ilmu ...)*, 2021.
- [10] A. Harist N, I. R. Munthe, and A. P. Juledi, "Implementasi Data Mining Algoritma Apriori untuk Meningkatkan Penjualan," *J. Tek. Inform. UNIKA St. Thomas*, 2021, doi: 10.54367/jtiust.v6i1.1276.