



Penerapan Data Mining *dalam* Menentukan Pola Penjualan

Alat Tulis Kantor di Kwek Photo Menggunakan Algoritma Apriori

Anggi Ervianti
Ibnu Rasyid Munthe, S.T., M.Kom
Angga Putra Juledi, S.Kom., M.Kom
Marnis Nasution., S.Kom., M.Kom



Penerapan Data Mining

dalam
Menentukan Pola Penjualan

Alat Tulis Kantor di Kwek Photo Menggunakan Algoritma Apriori

Anggi Erviyanti
Ibnu Rasyid Munthe, S.T., M.Kom
Angga Putra Juledi, S.Kom., M.Kom
Marnis Nasution., S.Kom., M.Kom



**Penerapan Data Mining dalam Menentukan
Pola Penjualan Alat Tulis Kantor di Kwek Photo
Menggunakan Algoritma Apriori**

Ditulis oleh:

**Anggi Ervianti
Ibnu Rasyid Munthe, S.T., M.Kom.
Angga Putra Juledi, S.Kom., M.Kom.
Marnis Nasution, S.Kom., M.Kom.**

Diterbitkan, dicetak, dan didistribusikan oleh

PT Literasi Nusantara Abadi Grup

Perumahan Puncak Joyo Agung Residence Blok B11 Merjosari

Kecamatan Lowokwaru Kota Malang 65144

Telp : +6285887254603, +6285841411519

Email: literasinusantaraofficial@gmail.com

Web: www.penerbitlitnus.co.id

Anggota IKAPI No. 340/JTI/2022



Hak Cipta dilindungi oleh undang-undang. Dilarang mengutip
atau memperbanyak baik sebagian ataupun keseluruhan isi buku
dengan cara apa pun tanpa izin tertulis dari penerbit.

Cetakan I, Agustus 2025

Perancang sampul: Rosyiful Aqli

Penata letak: D Gea Nuansa

ISBN : 978-634-234-400-2

viii + 78 hlm. ; 15,5x23 cm.

©Agustus 2025

Prakata

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga buku monograf berjudul *“Penerapan Data Mining dalam Menentukan Pola Penjualan Alat Tulis Kantor di Kwek Photo Menggunakan Algoritma Apriori”* ini dapat diselesaikan dengan baik. Buku ini disusun dengan tujuan memberikan gambaran komprehensif mengenai pemanfaatan teknologi data mining khususnya algoritma Apriori dalam mengidentifikasi pola-pola pembelian yang tersembunyi pada data transaksi penjualan alat tulis kantor di Kwek Photo.

Penerapan algoritma Apriori sebagai teknik utama dalam pengolahan data transaksi penjualan ini menjadi salah satu solusi inovatif dalam meningkatkan efektivitas pengambilan keputusan bisnis, terutama dalam hal pengelolaan stok, strategi pemasaran, dan pengembangan produk. Dengan memanfaatkan data mining, kami berharap buku ini tidak hanya memberikan wawasan teoretis, tetapi juga aplikatif bagi para praktisi bisnis, akademisi, serta mahasiswa yang tertarik mendalami bidang data mining dan analisis bisnis berbasis data.

Penulis menyadari bahwa perjalanan penyusunan buku ini tidak terlepas dari bantuan, dukungan, dan masukan banyak pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada seluruh pihak yang telah memberikan kontribusi baik secara langsung maupun tidak langsung, khususnya pemilik usaha Kwek Photo atas kesempatan dan data yang diberikan.

Akhir kata, penulis berharap buku ini dapat memberikan manfaat luas dan menjadi referensi yang bermanfaat dalam pengembangan riset dan praktik bisnis berbasis teknologi informasi. Kritik dan saran yang membangun sangat penulis harapkan demi peningkatan kualitas karya ini di masa mendatang.

Pengantar

Dalam era digital dan transformasi teknologi informasi yang pesat, pengolahan data besar (big data) telah menjadi aspek vital dalam pengambilan keputusan bisnis yang efektif dan efisien. Salah satu pendekatan yang semakin banyak diaplikasikan dalam dunia bisnis adalah data mining, yaitu teknik dan metode untuk menggali informasi tersembunyi dan pola-pola berharga dari kumpulan data yang sangat besar. Buku monograf ini hadir dengan tujuan memberikan pemahaman mendalam mengenai implementasi data mining, khususnya algoritma Apriori, dalam konteks menentukan pola penjualan alat tulis kantor di usaha ritel Kwek Photo.

Kajian ini difokuskan pada eksplorasi pola asosiasi antar produk yang dapat membantu manajemen dalam memahami perilaku konsumen serta dinamika pasar. Algoritma Apriori, sebagai salah satu metode asosiasi rule mining yang populer, dipilih karena kemampuannya dalam menemukan hubungan signifikan antar item yang sering muncul bersama dalam transaksi penjualan. Dengan penerapan algoritma ini, diharapkan diperoleh insight strategis untuk pengelolaan stok, penentuan promosi, dan peningkatan layanan yang berorientasi pada kebutuhan nyata pelanggan.

Selain aspek teknis, buku ini juga membahas secara komprehensif dampak strategis, tantangan yang dihadapi selama proses analisis data, serta rekomendasi praktis dan kebijakan yang relevan bagi pengembangan usaha Kwek Photo ke depan. Melalui pendekatan empiris dan berbasis data, diharapkan pembaca, terutama para praktisi dan akademisi di bidang manajemen bisnis dan teknologi informasi, dapat memperoleh pemahaman yang utuh dan aplikatif mengenai penerapan data mining dalam praktik nyata.

Akhir kata, buku ini bukan hanya berfungsi sebagai referensi akademis tetapi juga sebagai panduan praktis bagi pelaku usaha yang

ingin memanfaatkan kekuatan data untuk memperkuat daya saing dan keberlanjutan bisnis mereka dalam menghadapi tantangan pasar yang semakin kompetitif.

Daftar Isi

Prakata	iii
Pengantar	v
Daftar Isi	vii

BAB 1

Pendahuluan	1
Mengapa Pola Penjualan Penting untuk Bisnis Alat Tulis?	1
Membuka Pintu Data Mining untuk Analisis Bisnis	3
Fokus dan Batasan Penelitian	6
Tujuan dan Manfaat Penelitian	8
Gambaran Umum Buku	10

BAB 2

Landasan Teori	13
Menyelami Dunia Data Mining	13
Memahami Intisari Algoritma Apriori	15
Kelebihan dan Kekurangan Algoritma Apriori	17
Pilihan Teori Pendukung	17
Studi Sebelumnya sebagai Fondasi	21

BAB 3

Metode Penelitian dan Perancangan	23
Jenis Data dan Sumbernya	23
Persiapan dan Pembersihan Data	27
Merancang Strategi Analisis Data	34
Menentukan Parameter Kunci Analisis	36

BAB 4

Implementasi dan Pembahasan	39
Mengaplikasikan Algoritma Apriori Secara Manual	39
Pemanfaatan Software RapidMiner dalam Analisis	46
Langkah Teknis di RapidMiner	49
Hasil Analisis: Pola-pola Menarik yang Terungkap.....	52
Dampak Strategis Berdasarkan Temuan.....	61
Tantangan dan Solusi Selama Proses	63

BAB 5

Refleksi, Kesimpulan, dan Rekomendasi.....	67
Refleksi atas Proses Penelitian.....	67
Kesimpulan Utama yang Mendalam.....	69
Saran untuk Pengembangan Berikutnya	70
Tinjauan Manfaat Praktis dan Kebijakan	71
Daftar Pustaka.....	73
Riwayat Penulis.....	75

BAB 1

Pendahuluan

Mengapa Pola Penjualan Penting untuk Bisnis Alat Tulis?

Dalam konteks bisnis alat tulis kantor, pemahaman terhadap pola penjualan memegang peranan yang sangat strategis dan fundamental dalam upaya meningkatkan daya saing serta efektivitas pengambilan keputusan manajerial. Pola penjualan merupakan gambaran terstruktur dari kebiasaan membeli konsumen yang dapat mengungkap kecenderungan atau preferensi dalam memilih produk. Dengan mengenali pola-pola tersebut, sebuah toko retail dapat merumuskan strategi pemasaran yang tidak hanya sesuai dengan kebutuhan pasar tetapi juga mampu memaksimalkan keuntungan melalui pengelolaan produk yang lebih efisien.

Pentingnya pemahaman pola penjualan dalam bisnis alat tulis tidak lepas dari karakteristik pasar retail yang sangat dinamis dan kompetitif. Konsumen saat ini tidak hanya mencari produk secara individual, melainkan juga kombinasi produk yang saling melengkapi atau memang biasa dibeli secara bersamaan. Sebagai contoh, seorang konsumen yang



membeli kertas A4 mungkin juga cenderung membeli tinta printer atau alat tulis lainnya dalam transaksi yang sama. Informasi ini sangat berharga untuk merancang tawaran bundling produk, promosi silang, ataupun tata letak yang strategis di dalam toko agar bisa meningkatkan volume penjualan.

Di sisi lain, volume data transaksi yang besar dan kompleksitas pembelian yang variatif sering kali menjadi kendala dalam mendeteksi pola penjualan secara manual. Tanpa dukungan teknik analisis yang memadai, peluang untuk menggali wawasan penting dari data yang ada menjadi sangat terbatas. Oleh karena itu, penerapan teknik data mining, khususnya algoritma Apriori, menjadi solusi yang relevan. Algoritma ini dirancang untuk menemukan aturan asosiasi kombinasi produk yang paling sering muncul bersama dalam suatu transaksi. Melalui metode ini, bisnis dapat mengidentifikasi frequent itemsets yang kritical sebagai dasar pengambilan keputusan pemasaran.

Selain itu, pemahaman yang baik terhadap pola penjualan juga memungkinkan efisiensi pengelolaan inventaris. Dengan mengetahui produk mana yang cenderung dibeli bersama, toko dapat memperkirakan kebutuhan stok secara lebih akurat dan menghindari kelebihan atau kekurangan persediaan yang merugikan. Optimalisasi stok ini tidak hanya mengurangi biaya penyimpanan tetapi juga menjaga kesinambungan layanan terhadap pelanggan.

Tidak kalah penting adalah aspek strategis pemasaran yang dapat dikembangkan berdasarkan pola penjualan yang dianalisis. Dengan aturan asosiasi yang terungkap, toko bisa mengimplementasikan teknik penjualan seperti rekomendasi produk, program loyalitas, atau diskon khusus untuk paket produk tertentu, sehingga mampu meningkatkan engagement pelanggan dan memperkuat posisi bisnis di pasar. Secara tidak langsung, pola-pola tersebut berfungsi sebagai peta jalan dalam memahami perilaku konsumen yang dinamis.

Penelitian yang telah dilakukan pada data transaksi penjualan alat tulis kantor di Kwek Photo menunjukkan bahwa penerapan algoritma Apriori

mampu menghasilkan pola pembelian yang signifikan. Penggunaan parameter support dan confidence secara tepat dapat menyaring aturan asosiasi yang relevan dan aplikatif bagi toko retail tersebut. Temuan ini memperkuat pentingnya pendekatan berbasis data mining agar pengambilan kebijakan pemasaran dan penataan produk dapat dilakukan secara ilmiah dan sistematis.

Secara menyeluruh, dari perspektif manajemen bisnis, kemampuan untuk mengidentifikasi dan memanfaatkan pola penjualan adalah kunci untuk memenangkan persaingan di pasar retail alat tulis yang penuh tantangan. Pendekatan ini tidak hanya memperkaya wawasan manajerial tetapi juga menambah nilai kompetitif organisasi melalui adaptasi cepat terhadap perubahan perilaku konsumen. Oleh karena itu, pengembangan sistem analisis pola penjualan berbasis data mining perlu diintegrasikan sebagai bagian dari strategi bisnis jangka panjang demi mendukung pertumbuhan dan keberlanjutan usaha.

Membuka Pintu Data Mining untuk Analisis Bisnis

Dalam era digital dan persaingan bisnis yang semakin ketat, data telah menjadi aset strategis yang memegang peranan penting dalam pengambilan keputusan yang efektif dan efisien. Khususnya dalam bisnis penjualan alat tulis kantor, volume data transaksi yang besar dan kompleks menuntut pendekatan inovatif dalam pengelolaan dan penafsiran data. Data mining hadir sebagai solusi inovatif yang membuka pintu baru untuk menggali wawasan dan pola tersembunyi dari data transaksi yang sebelumnya sulit diakses secara manual. Dengan penerapan teknik data mining, bisnis dapat mengekstrak informasi yang bernilai dari tumpukan data mentah menjadi pengetahuan yang dapat digunakan untuk meningkatkan kinerja operasional dan strategi pemasaran.

Data mining adalah proses otomatis atau semi-otomatis yang mengintegrasikan metode statistik, kecerdasan buatan, dan algoritma

khusus untuk menemukan pola, tren, dan hubungan signifikan dalam kumpulan data besar. Dalam konteks bisnis alat tulis, data mining memungkinkan pemilik usaha untuk memahami perilaku pembelian konsumen secara mendalam, termasuk kombinasi produk yang sering dibeli bersamaan, frekuensi pembelian, hingga pengelompokan pelanggan berdasarkan karakteristik tertentu. Melalui proses ini, manajemen dapat merumuskan strategi yang lebih tepat sasaran dan berbasis evidence, mengurangi ketidakpastian dalam pengambilan keputusan.

Salah satu algoritma data mining yang populer dalam menganalisis pola penjualan adalah algoritma Apriori. Algoritma ini secara khusus digunakan untuk menemukan aturan asosiasi dalam data transaksi—misalnya, untuk mengetahui produk mana yang sering dibeli dalam satu keranjang belanja. Dengan menghitung frequent itemsets berdasarkan parameter support (tingkat kemunculan kombinasi produk) dan confidence (kekuatan aturan asosiasi), Apriori membantu mengidentifikasi hubungan yang memiliki signifikansi statistik dan relevansi bisnis. Penerapan algoritma ini dalam analisis penjualan alat tulis di Kwek Photo membuktikan efektivitasnya dalam memberikan insight strategis yang sebelumnya tersembunyi.

Manfaat utama dari data mining dalam analisis bisnis tidak hanya terbatas pada pemahaman pola penjualan, melainkan juga merambah ke optimalisasi pengelolaan inventaris. Dengan mengetahui pola pembelian yang sering terjadi, bisnis dapat mengelola stok barang secara lebih akurat sehingga meminimalkan biaya pemborosan akibat kelebihan stok atau kehilangan potensi penjualan karena kehabisan barang. Selain itu, data mining mendukung pengembangan strategi pemasaran yang lebih personal dan dinamis, seperti penawaran bundling, promosi silang, dan rekomendasi produk yang sesuai dengan perilaku konsumen.

Lebih jauh, integrasi data mining ke dalam sistem informasi bisnis membuka peluang untuk inovasi berkelanjutan. Dengan fasilitas pengumpulan data yang terus berlangsung dan kemampuan analitik yang adaptif, bisnis dapat memantau perubahan tren dan preferensi pasar secara real-time atau periodik. Hal ini memungkinkan mereka untuk segera

menyesuaikan strategi pemasaran atau produk sesuai dengan kebutuhan pasar yang berkembang, mempertahankan keunggulan kompetitif dalam menghadapi persaingan.

Dalam perspektif akademik, pengaplikasian data mining dalam analisis penjualan alat tulis menawarkan kontribusi penting bagi bidang ilmu manajemen dan teknologi informasi. Studi yang sistematis menggunakan algoritma Apriori membuka jalur untuk penelitian lanjutan yang dapat mengembangkan metode analisis yang lebih canggih atau menggabungkan teknik lain seperti clustering dan classification. Selain itu, hasil penelitian ini memberikan dasar empiris bagi pengembangan sistem pendukung keputusan (decision support system) yang lebih efektif di lingkup bisnis retail kecil hingga menengah.

Penelitian ini berdasarkan data riil dari Kwek Photo mengilustrasikan bagaimana transformasi data menjadi informasi yang bermakna dapat terjadi melalui penerapan data mining. Dengan dukungan perangkat lunak analitik dan metode yang tepat, pola penjualan yang sebelumnya tersembunyi di balik jutaan transaksi kini terkuak dengan jelas. Hal ini membuka kesempatan untuk meningkatkan profitabilitas dan efisiensi usaha, sekaligus memperkuat loyalitas pelanggan melalui pemahaman yang lebih baik terhadap kebutuhan mereka.

Secara keseluruhan, data mining bukan sekadar alat teknis, melainkan suatu inovasi revolusioner yang mengubah cara bisnis mengelola data dan mengambil keputusan strategis. Dalam bisnis alat tulis, khususnya di pasar ritel yang kompetitif, pemanfaatan teknologi ini menjadi sangat krusial untuk meningkatkan daya saing dan mengoptimalkan sumber daya yang ada. Oleh karena itu, membuka pintu data mining dalam analisis bisnis merupakan langkah fundamental menuju transformasi digital yang berkelanjutan dan sukses di masa depan.

Fokus dan Batasan Penelitian

Dalam sebuah penelitian, penentuan fokus dan batasan penelitian memegang peranan krusial guna membantu pembaca memahami secara jelas cakupan kajian yang dilakukan serta agar penelitian tetap berada dalam koridor yang terarah dan terkendali. Fokus penelitian merujuk pada aspek utama yang menjadi titik perhatian dan kajian intensif dalam penelitian tersebut, sedangkan batasan penelitian adalah parameter yang mendefinisikan ruang lingkup dan kekhususan dari studi yang dijalankan. Penjelasan yang rinci mengenai fokus dan batasan ini tidak hanya membantu mengarahkan proses penelitian tetapi juga menghindarkan dari perluasan topik yang tidak relevan yang dapat melemahkan kualitas dan ketajaman analisis.

Pada penelitian ini, fokus utama diarahkan pada penerapan algoritma Apriori dalam menganalisis pola penjualan produk alat tulis kantor di Kwek Photo. Penjelasan fokus penelitian ini secara eksplisit mengedepankan tujuan untuk memahami perilaku pembelian konsumen melalui data transaksi yang tersedia. Dalam hal ini, algoritma Apriori dipilih sebagai metode data mining yang efektif untuk mengidentifikasi asosiasi dan pola berulang antar produk yang dibeli bersamaan. Konsentrasi pada algoritma ini memberikan bingkai yang jelas dalam menjalankan analisis, yakni menguji parameter support dan confidence sebagai dasar evaluasi aturan asosiasi yang terbentuk.

Batasan penelitian kemudian mendefinisikan ruang lingkup agar kajian tetap terfokus dan dapat diukur keberhasilannya. Batasan yang diterapkan dalam penelitian ini meliputi aspek temporal data transaksi penjualan yang dianalisis, yaitu periode tertentu yang telah ditetapkan sesuai dengan ketersediaan data di Kwek Photo. Selain itu, penelitian ini membatasi cakupan hanya pada produk alat tulis kantor, sehingga efek atau pola dari produk lain di luar kategori tersebut tidak menjadi bagian dari analisis dan hasil penelitian. Batasan juga mencakup penggunaan satu algoritma spesifik, yakni Apriori, tanpa eksplorasi metode data mining lain seperti clustering atau classification untuk menjaga konsistensi analisis.

Penetapan fokus dan batasan ini sangat penting untuk memberikan kepastian metodologis dan menghindari generalisasi yang berlebihan terhadap hasil penelitian. Dengan fokus pada pola pembelian dalam konteks alat tulis kantor dan penerapan algoritma Apriori, penelitian dapat menggali insight yang mendalam dan relevan bagi pengembangan strategi pemasaran dan pengelolaan inventaris di Kwek Photo. Hal ini pula menghindarkan dari bias yang mungkin muncul jika ruang lingkup terlalu luas atau beragam metode digunakan tanpa konsistensi.

Batasan yang jelas juga membantu dalam evaluasi hasil penelitian secara ilmiah. Pembaca dapat mengetahui dengan tepat konteks data dan metode yang digunakan serta memahami kondisi dan asumsi yang membatasi hasil yang dicapai. Dengan demikian, temuan penelitian dapat dipertanggungjawabkan dan dijadikan dasar untuk rekomendasi yang spesifik dan aplikatif. Misalnya, rekomendasi yang dihasilkan fokus pada optimasi penjualan alat tulis di Kwek Photo dan tidak secara langsung menysasar sektor usaha lain yang berbeda konteks dan karakteristik pasar.

Dari perspektif pengelolaan penelitian, fokus dan batasan ini juga berfungsi sebagai pedoman pelaksanaan serta pengendalian kegiatan penelitian. Para peneliti dapat menjaga agar analisis dan interpretasi data tetap relevan dan terarah sesuai tujuan awal. Dengan demikian, efisiensi waktu dan sumber daya dapat lebih terjaga karena tidak terbuang untuk wilayah kajian yang tidak prioritas atau tidak memberikan kontribusi signifikan terhadap tujuan penelitian.

Lebih lanjut, fokus dan batasan yang jelas mendukung aspek replikasi dan pengembangan penelitian. Penelitian lain yang ingin melakukan studi serupa dapat menggunakan cakupan yang telah ditetapkan ini sebagai titik acuan. Jika pengembangannya ingin memperluas ruang lingkup, misalnya memasukkan produk lain atau algoritma data mining tambahan, maka perbedaan ini dapat ditegaskan secara metodologis sehingga menambah pengetahuan yang saling melengkapi di bidang kajian yang sama.

Secara keseluruhan, penguraian fokus dan batasan penelitian dalam studi penerapan algoritma Apriori untuk pola penjualan alat

tulis kantor di Kwek Photo membentuk fondasi metodologis yang kokoh. Dengan pendekatan yang terstruktur dan terarah, penelitian mampu memberikan kontribusi empiris yang valid dan terpercaya, serta mendukung pengambilan keputusan bisnis yang berbasis data. Kejelasan ruang lingkup ini juga memudahkan pembaca untuk mengikuti logika dan alur penelitian sehingga menumbuhkan kepercayaan terhadap hasil dan rekomendasi yang disajikan.

Tujuan dan Manfaat Penelitian

Penelitian mengenai penerapan data mining untuk menentukan pola penjualan alat tulis kantor merupakan upaya strategis dalam memanfaatkan data transaksi yang selama ini banyak terkumpul namun kurang diolah secara optimal. Tujuan utama dari penelitian ini adalah untuk mengidentifikasi pola pembelian konsumen secara sistematis menggunakan algoritma Apriori, yang memungkinkan untuk menemukan asosiasi produk yang sering dibeli bersama dan memberikan insight untuk perencanaan strategi pemasaran yang lebih efektif. Dengan demikian, penelitian ini tidak hanya menghasilkan pola data statistik tetapi juga memberikan kontribusi konkret dalam meningkatkan pengelolaan bisnis retail, khususnya di sektor alat tulis kantor pada toko Kwek Photo.

Penentuan tujuan dalam penelitian ini sangat penting karena mengarahkan seluruh proses metodologis, mulai dari pengumpulan data, pemilihan metode, hingga analisis dan interpretasi hasil. Fokus pada aplikasi algoritma Apriori membuka peluang untuk menggali hubungan kompleks antar item penjualan yang tersembunyi dalam kumpulan besar data transaksi, yang secara manual sulit sekali ditemukan. Parameter penting seperti nilai support dan confidence diuji untuk mendapatkan pengaturan paling optimal yang mendukung validitas dan relevansi pola yang ditemukan. Tujuan ini menegaskan bahwa hasil penelitian diharapkan tidak hanya akademis, tetapi juga aplikatif untuk mendukung pengambilan keputusan bisnis berbasis data (data-driven decision making).

Manfaat praktis dari penelitian ini sangat beragam dan menysasar berbagai aspek operasional dan manajerial dalam bisnis alat tulis kantor. Pertama, pemahaman lebih baik tentang perilaku konsumen memungkinkan pengelola toko melakukan segmentasi produk yang lebih tepat, misalnya dengan mengelompokkan produk yang sering dibeli bersamaan agar ditempatkan berdekatan, sehingga meningkatkan peluang cross-selling dan bundling produk. Kedua, hasil penelitian dapat digunakan untuk merancang strategi pemasaran yang lebih terukur dan efektif, seperti diskon pada produk komplementer atau kampanye promosi yang disesuaikan dengan pola pembelian pelanggan.

Selain itu, penelitian juga berkontribusi pada peningkatan efisiensi pengelolaan persediaan dan inventaris. Dengan pola penjualan yang jelas, pengecer dapat melakukan prediksi kebutuhan stok secara akurat, mengurangi risiko kelebihan stok atau kekurangan barang, yang secara langsung berdampak pada pengelolaan modal dan operasional yang lebih efisien. Di sisi lain, penelitian ini menyediakan kerangka evaluasi yang berguna untuk memantau perubahan perilaku konsumen dari waktu ke waktu, sehingga strategi yang dihasilkan dapat terus diperbaharui agar tetap relevan dengan dinamika pasar.

Manfaat tersebut tidak hanya dirasakan oleh pengelola toko, tetapi juga oleh pemangku kepentingan lain di sektor bisnis retail, terutama yang ingin mengimplementasikan teknologi data mining dalam aktivitas sehari-hari. Penelitian ini memberikan dasar empiris dan metodologis yang dapat menjadi acuan bagi studi selanjutnya maupun pengembangan sistem aplikasi berbasis data mining yang lebih canggih dan terintegrasi dengan sistem bisnis modern. Secara teoretis, penelitian ini memperluas literatur terkait penggunaan algoritma apriori di bidang manajemen pemasaran dan strategi bisnis dengan konteks industri alat tulis kantor yang spesifik.

Selanjutnya, manfaat jangka panjangnya juga mencakup peningkatan daya saing toko Kwek Photo di pasar yang semakin ketat. Dengan data dan pola yang dihasilkan, perusahaan dapat membuat keputusan lebih cepat, akurat, dan berbasis bukti daripada hanya mengandalkan intuisi

atau pengalaman semata. Hal ini menjadi nilai tambah dalam transformasi digital yang sedang berlangsung di semua lini bisnis, mendukung keberlanjutan usaha melalui inovasi dan peningkatan kualitas layanan pelanggan.

Melalui penelitian ini, harapannya dunia usaha, khususnya pengusaha retail alat tulis kantor, dapat melihat nilai strategis dari implementasi data mining sebagai alat bantu penting dalam mengelola bisnis yang lebih adaptif dan responsif terhadap perubahan perilaku konsumen. Selain itu, penelitian ini juga membuka ruang bagi pengembangan teknologi baru dan metode analisis data yang lebih kompleks seperti machine learning dan artificial intelligence untuk mengoptimalkan fungsi bisnis secara keseluruhan.

Gambaran Umum Buku

Buku ini disusun dengan tujuan memberikan pemahaman mendalam tentang penerapan data mining dalam menentukan pola penjualan alat tulis kantor, khususnya di toko retail seperti Kwek Photo. Melalui pendekatan sistematis dan metodologis, pembaca diajak untuk memahami berbagai konsep dasar, teknik analisis data, serta implementasi algoritma Apriori yang digunakan sebagai inti dari studi ini. Dengan memberikan gambaran umum di awal buku, diharapkan pembaca dapat mengikuti perkembangan isi buku secara bertahap dan memiliki kerangka pemikiran yang jelas sejak awal.

Bagian pembuka buku ini berperan sebagai pintu gerbang yang memperkenalkan topik utama serta struktur keseluruhan yang akan dibahas. Di sini, pembaca disiapkan untuk menyelami konteks permasalahan bisnis dalam penjualan alat tulis, mengapa analisis pola penjualan sangat penting, dan bagaimana teknologi data mining dapat menjadi solusi inovatif yang mengoptimalkan pengambilan keputusan bisnis. Penekanan pada konteks bisnis nyata memberikan nilai aplikatif

sehingga isi buku tidak hanya bersifat teoritis, tetapi juga pragmatis dan relevan dengan kebutuhan praktis.

Penjelasan dalam bagian gambaran umum buku mencakup uraian terkait latar belakang penelitian, tujuan, serta manfaat yang ingin dicapai, yang kemudian dilanjutkan dengan pemaparan metodologi yang akan digunakan. Hal ini membangun fondasi pemahaman bagi pembaca untuk menyambut bab-bab berikutnya yang lebih detail membahas aspek teknis dan analisis data. Secara sistematis, buku ini membagi isi ke dalam beberapa bab utama yang saling berkesinambungan, mulai dari pendahuluan, tinjauan pustaka, metodologi penelitian, analisis data hingga penarikan kesimpulan dan rekomendasi bagi dunia usaha.

Selain itu, gambaran umum buku juga memberikan penjelasan tentang struktur penyajian isi agar pembaca dapat mengantisipasi apa saja yang akan dipelajari dan bagaimana tiap bab terintegrasi dalam membangun argumen yang kuat. Bab-bab yang disusun mengikuti alur logis dimulai dari pemahaman konsep dasar data mining, pengenalan khusus terhadap algoritma Apriori, aplikasi dalam konteks penjualan alat tulis, hingga evaluasi hasil analisis yang memberikan insight bisnis penting. Dengan demikian, buku ini tidak hanya sebagai bahan bacaan, tetapi juga sebagai panduan praktis bagi pembaca yang ingin mengimplementasikan data mining dalam bisnis mereka.

Untuk mendukung pemahaman, buku ini juga dilengkapi dengan berbagai tabel, grafik, dan contoh kasus yang memvisualisasikan data dan hasil penelitian secara jelas. Elemen-elemen visual ini diintegrasikan secara strategis agar pembaca dapat lebih mudah menangkap makna setiap analisis dan interpretasi yang disampaikan. Pendekatan ini sangat penting mengingat kompleksitas data mining dan perlunya pemahaman visualisasi hasil agar tidak hanya berfokus pada teori, tetapi juga aplikasi nyata dan manfaat yang dapat dirasakan oleh pelaku bisnis.

Gambaran umum ini juga menyoroti relevansi buku bagi berbagai kalangan, mulai dari akademisi, praktisi bisnis, hingga mahasiswa yang ingin mendapatkan wawasan terbaru tentang teknologi data mining

dalam bidang manajemen pemasaran. Dengan bahasa yang formal namun komunikatif, isi buku diformat agar dapat diakses oleh pembaca dengan latar belakang yang beragam. Penjelasan bertahap dengan contoh nyata dari data penjualan alat tulis kantor memperkuat daya tarik dan nilai praktis buku ini.

Di sisi lain, bagian ini mengarahkan pembaca untuk memahami bahwa buku ini merupakan hasil dari penelitian yang mendalam dan valid secara ilmiah, mengacu pada berbagai sumber terpercaya dan literatur terkini. Hal ini meningkatkan kredibilitas isi buku sebagai referensi akademik yang dapat dijadikan dasar untuk penelitian lanjutan atau pengembangan sistem berbasis data mining di masa depan.

Dengan gambaran umum yang kuat dan terstruktur ini, diharapkan pembaca merasa tertantang dan termotivasi untuk menggali lebih dalam bab-bab berikutnya. Pendekatan bertahap yang disajikan secara sistematis memberikan jalur belajar yang jelas dan konsisten sehingga baik pembaca pemula maupun yang telah berpengalaman dapat memperoleh manfaat maksimal dari isi buku.

BAB 2

Landasan Teori

Menyelami Dunia Data Mining

Data mining adalah proses sistematis yang bertujuan mengekstraksi informasi atau pola yang bermanfaat dari dataset berukuran besar melalui penerapan teknik statistik, pembelajaran mesin, dan kecerdasan buatan [1]. Data mining sering disebut sebagai bagian integral dari proses Knowledge Discovery in Databases (KDD), yang mencakup tahapan pengumpulan data, prapemrosesan, analisis pola, dan interpretasi hasil untuk menghasilkan pengetahuan yang dapat mendukung pengambilan keputusan strategis [2][3]

Proses data mining melibatkan beberapa teknik utama, seperti: Klasifikasi, yang mengelompokkan data ke dalam kategori tertentu berdasarkan atribut yang relevan ([4]. Klasterisasi, yang mengelompokkan data berdasarkan kesamaan tanpa kategori yang telah ditentukan sebelumnya Asosiasi, yang digunakan untuk mengidentifikasi hubungan antar item dalam data, seperti pola pembelian di sektor ritel [5] Selain itu, penerapan data mining tidak hanya terbatas pada sektor bisnis, tetapi juga pada pendidikan untuk analisis perilaku siswa, di bidang kesehatan untuk

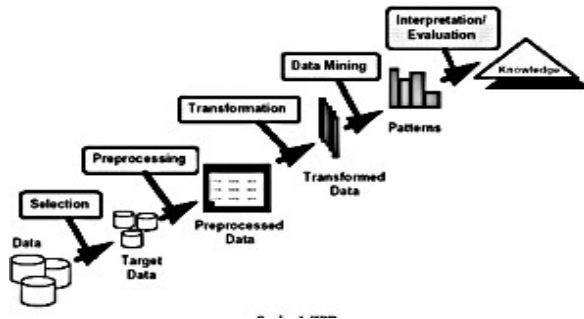


prediksi hasil perawatan pasien, dan dalam pengelolaan kota pintar untuk efisiensi layanan publik [6][7]

Proses dasar dalam data mining terdiri dari beberapa langkah utama yang harus dijalani secara berurutan dan terstruktur. Pertama adalah pengumpulan data yang sudah diolah atau mentah dari berbagai sumber. Setelah data terkumpul, dilakukan tahap pembersihan data (data cleansing) untuk menghilangkan data yang tidak valid, duplikat, atau tidak relevan. Kemudian, proses preprocessing dilakukan untuk mengkonversi dan mentransformasikan data menjadi bentuk yang sesuai untuk dianalisis, seperti normalisasi atau pengisian nilai yang hilang.

Tahapan KDD:

1. Seleksi Data (Data Selection): Proses memilih data yang relevan untuk dianalisis dari kumpulan data yang tersedia.
2. Prapemrosesan Data (Data Preprocessing): Membersihkan dan mempersiapkan data dengan mengatasi inkonsistensi, menangani nilai yang hilang, atau menghapus data yang tidak relevan.
3. Transformasi Data (Data Transformation): Mengubah data ke dalam format yang sesuai untuk analisis lebih lanjut, seperti normalisasi atau agregasi.
4. Penambangan Data (Data Mining): Penerapan algoritma dan teknik analitik untuk menemukan pola atau hubungan dalam data.
5. Evaluasi dan Interpretasi (Evaluation and Interpretation): Menganalisis dan memvalidasi pola yang ditemukan untuk memastikan relevansi dan akurasi terhadap tujuan penelitian atau bisnis.
6. Penyajian Pengetahuan (Knowledge Representation): Menyajikan hasil dalam bentuk yang mudah dimengerti, seperti laporan, grafik, atau visualisasi.



Gambar 2.1 Proses Data Mining dalam Kerangka KDD

Memahami Intisari Algoritma Apriori

Algoritma Apriori adalah salah satu teknik utama dalam data mining yang bertujuan untuk menemukan pola asosiatif atau aturan hubungan antara item-item dalam kumpulan data transaksi secara efisien. Algoritma ini sangat populer dan banyak digunakan dalam bidang analisis pasar dan bisnis untuk mengidentifikasi pola pembelian tersembunyi yang sering terjadi bersamaan dalam transaksi pelanggan. Dalam konteks penelitian ini, algoritma Apriori digunakan untuk mengungkap pola pembelian alat tulis kantor, sehingga membantu pemilik bisnis dalam mengambil keputusan strategis yang lebih tepat berdasarkan pola tersebut.

Inti dari algoritma Apriori adalah menemukan itemset frekuent yang memenuhi batas minimum support yang telah ditentukan sebelumnya. Support adalah ukuran frekuensi kemunculan suatu kombinasi item dalam data transaksi. Algoritma ini bekerja berdasarkan prinsip dasar bahwa semua subset dari itemset frekuent juga harus frekuent. Artinya, jika suatu kombinasi item tidak memenuhi batas minimum support, maka semua kombinasi yang mengandung item tersebut tidak perlu dihitung lebih lanjut karena dipastikan tidak akan memenuhi syarat.

Proses kerja algoritma Apriori dimulai dengan mengidentifikasi item tunggal yang sering muncul dalam data transaksi, kemudian menggabungkan item-item tersebut menjadi kombinasi dua item, lalu tiga item, dan seterusnya, dengan mengevaluasi masing-masing kombinasi

berdasarkan nilai support. Kombinasi yang gagal memenuhi nilai support yang diharapkan akan segera dihilangkan dari perhitungan selanjutnya, sehingga mengurangi kompleksitas dan beban komputasi. Dengan demikian, algoritma ini mengaplikasikan teknik pruning yang sangat efisien dalam mengeliminasi kandidat yang tidak relevan.

Selain support, algoritma Apriori juga mengukur kekuatan aturan asosiasi menggunakan nilai confidence, yaitu probabilitas terjadinya item B ketika item A sudah terjadi di dalam satu transaksi yang sama (rule: $A \rightarrow B$). Confidence yang tinggi menunjukkan bahwa aturan tersebut dapat diandalkan untuk prediksi atau rekomendasi, seperti bundling produk atau penataan strategis dalam display penjualan. Contohnya, jika aturan “pembelian pena \rightarrow pembelian buku catatan” memiliki confidence tinggi, maka penjual bisa memanfaatkan informasi ini untuk strategi pemasaran.

Dalam konteks bisnis, algoritma Apriori sangat berguna dalam mengidentifikasi pola pembelian tersembunyi yang tidak terlihat secara langsung dari data transaksi mentah. Misalnya, tabel hasil penelitian yang ditemukan dalam naskah Anda menunjukkan bahwa seringkali pelanggan membeli pena dan buku catatan secara bersamaan dengan tingkat support cukup tinggi dan confidence yang memadai. Penemuan pola ini membantu bisnis merancang paket produk yang lebih menarik dan efisien, serta mengelola persediaan dengan lebih baik untuk memenuhi permintaan pelanggan.

Algoritma Apriori juga dapat menyederhanakan pemahaman perilaku konsumen dengan menganalisis kumpulan data transaksi besar secara otomatis dan sistematis. Dengan kemampuannya dalam menghasilkan aturan asosiasi yang kuat, algoritma ini bermanfaat sebagai fondasi dalam sistem rekomendasi dan pengambilan keputusan berbasis data. Namun, algoritma ini memiliki keterbatasan, terutama dalam hal performa ketika dataset sangat besar dan kompleks, sehingga penting untuk melakukan optimasi atau menggabungkannya dengan teknik lain.

Kelebihan dan Kekurangan Algoritma Apriori

Kelebihan

1. Sederhana dan Mudah Diimplementasikan: Algoritma Apriori dikenal dengan struktur yang sederhana, menjadikannya mudah untuk dipahami dan diterapkan dalam berbagai konteks. Keunggulan ini menjadikannya pilihan utama dalam berbagai aplikasi analisis data, terutama dalam riset yang melibatkan pola pembelian atau hubungan antar produk.
2. Identifikasi Pola Asosiasi: Algoritma ini efektif dalam menemukan hubungan atau asosiasi antar item dalam data transaksi. Hal ini sangat berguna dalam merancang strategi pemasaran yang tepat, seperti promosi bundling produk yang sering dibeli bersamaan.

Kekurangan

1. Kompleksitas Waktu: Salah satu kekurangan signifikan dari algoritma Apriori adalah kebutuhan waktu komputasi yang tinggi, terutama pada dataset yang sangat besar. Hal ini dapat mempengaruhi efisiensi dalam penerapan algoritma pada data transaksi yang kompleks dan besar [8]
2. Masalah Kombinatorial: Algoritma Apriori cenderung menghadapi masalah kombinatorial, yaitu meningkatnya jumlah kombinasi itemset yang mungkin secara eksponensial. Ini bisa menghambat efisiensi algoritma dalam pengolahan data yang sangat besar dan kompleks ([9])

Pilihan Teori Pendukung

Penelitian oleh [10] Penelitian ini bertujuan untuk menganalisis pola pembelian produk di sebuah toko retail menggunakan algoritma Apriori. Fokus penelitian adalah untuk mengidentifikasi pola pembelian produk yang sering dibeli bersamaan dan menghasilkan aturan asosiasi yang dapat digunakan dalam strategi pemasaran. Hasil penelitian menunjukkan

bahwa algoritma Apriori mampu mengidentifikasi kombinasi produk yang relevan untuk promosi bundling.

1. Hubungan: Penelitian ini memberikan dasar teoritis untuk penggunaan algoritma Apriori dalam analisis data retail.
2. Kekurangan: Penelitian ini tidak membahas pengaruh variasi parameter support dan confidence terhadap hasil analisis.

Penelitian oleh [11] Penelitian ini berfokus pada penggunaan data mining untuk menemukan pola pembelian produk di sektor retail elektronik. Algoritma Apriori digunakan untuk mengevaluasi data transaksi dan menghasilkan rekomendasi bundling produk. Hasil penelitian menunjukkan bahwa penggunaan algoritma ini dapat meningkatkan efisiensi strategi pemasaran.

1. Hubungan: Penelitian ini menunjukkan potensi algoritma Apriori dalam meningkatkan efisiensi pemasaran, yang relevan dengan konteks penelitian di Kwek Photo.
2. Kekurangan: Penelitian ini tidak menjelaskan detail proses preprocessing data yang digunakan.

Penelitian oleh Penelitian ini mengkaji penerapan algoritma Apriori untuk analisis pola pembelian produk kebutuhan sehari-hari di sebuah supermarket. Fokus penelitian adalah pada pengoptimalan tata letak produk berdasarkan pola pembelian yang sering terjadi. Hasilnya menunjukkan bahwa tata letak yang dioptimalkan berdasarkan hasil analisis dapat meningkatkan pengalaman pelanggan dan penjualan toko.

1. Hubungan: Temuan ini relevan untuk mendukung strategi penataan produk di Kwek Photo.
2. Kekurangan: Penelitian ini tidak memperhitungkan perubahan preferensi pelanggan seiring waktu.

Dengan pemahaman mendalam mengenai algoritma Apriori ini, pembaca diharapkan dapat mengapresiasi prinsip kerja, keunggulan, dan aplikasi praktisnya dalam dunia bisnis. Bab ini mengedepankan aspek teoretis sekaligus memberikan wawasan aplikatif yang memadai

untuk analisis data transaksi dengan tujuan menemukan pola pembelian tersembunyi yang dapat diaplikasikan sebagai dasar pembuatan keputusan yang lebih strategis dan efektif.

Tabel 1 Penelitian Terdahulu

Judul Penelitian	Penulis	Hasil Penelitian	Hubungan dengan Penelitian Anda	Kekurangan Penelitian
Implementasi Data Mining Dengan Metode Apriori dalam Menentukan Pola Pemilihan Pemeriksaan Kimia [12]	D. Romdon, I. Kholil	Menemukan pola asosiatif untuk meningkatkan efisiensi layanan pemeriksaan kimia melalui analisis data transaksi.	Relevan dalam teknik penerapan algoritma Apriori untuk mengidentifikasi pola dalam dataset besar.	Fokus pada pemeriksaan kimia, sehingga tidak membahas konteks retail atau parameter yang relevan untuk penjualan produk.
Penerapan Algoritma Apriori dalam Menentukan Pola Penjualan Produk Pertanian [13]	W.H. Sirait, S. Saifulallah, J. Jalaludin	Mengidentifikasi frequent itemsets dalam data penjualan produk pertanian untuk rekomendasi tata letak dan promosi.	Memberikan wawasan tentang penerapan algoritma Apriori untuk pola penjualan produk fisik.	Parameter algoritma yang digunakan terbatas, tanpa memperhitungkan variabel tambahan seperti tren musiman.

Judul Penelitian	Penulis	Hasil Penelitian	Hubungan dengan Penelitian Anda	Kekurangan Penelitian
Penggunaan Algoritma Apriori untuk Menemukan Pola Belanja Konsumen Sebagai Promo Produk [2023]	Fakultas Teknologi Informasi UM Pasuruan	Menghasilkan pola pembelian untuk membantu perancangan promo berbasis data belanja konsumen.	Sejalan dengan penelitian Anda yang berfokus pada analisis data transaksi untuk meningkatkan strategi pemasaran.	Tidak mendalam dalam mengevaluasi pengaruh parameter algoritma pada efektivitas hasil yang dihasilkan.
Analisa Pola Penjualan Makanan Dengan Penerapan Algoritma Apriori [2021]	F. Handayanna, E. Irfiani	Algoritma Apriori digunakan untuk menentukan pola pembelian makanan sehingga membantu penyusunan strategi penempatan produk.	Memberikan contoh konkret implementasi algoritma Apriori dalam sektor retail untuk pola penjualan.	Hanya berfokus pada data makanan, tidak membahas kasus kategori produk lain seperti alat tulis kantor.
Perbandingan Algoritma Apriori dan FP-Growth dalam Menentukan Pola Penjualan Pupuk [2024]	Jurnal RESISTOR	Menganalisis efektivitas kedua algoritma dalam mengidentifikasi pola penjualan pupuk berdasarkan kecepatan dan kualitas hasil.	Membandingkan alternatif algoritma untuk meningkatkan efisiensi dalam analisis pola penjualan.	Lebih fokus pada perbandingan algoritma daripada memperdalam aplikasi khusus algoritma Apriori pada kasus retail spesifik.

Tabel ini secara sistematis menguraikan judul penelitian, penulis, hasil penelitian, hubungan dengan penelitian yang sedang dilakukan, serta kekurangan dari masing-masing penelitian sebelumnya. Tabel ini memperlihatkan bagaimana penelitian sebelumnya telah menggunakan algoritma Apriori untuk menemukan pola pembelian dalam berbagai

bidang seperti pemeriksaan kimia, produk pertanian, dan pola belanja secara umum. Masing-masing penelitian tersebut memiliki relevansi dengan fokus penelitian Anda yang juga menggunakan algoritma Apriori untuk menganalisis pola pembelian alat tulis kantor.

Analisis tabel ini membantu pembaca memahami bahwa walaupun algoritma Apriori telah banyak diterapkan, setiap penelitian memiliki konteks dan parameter yang berbeda, serta ada ruang untuk pengembangan lebih lanjut. Ini menegaskan pentingnya penelitian Anda dalam mengisi gap dengan mengkaji pola pembelian dalam konteks yang lebih spesifik dan variabel parameter yang lebih komprehensif. Dengan demikian, tabel ini memperkuat argumen penggunaan frequent itemsets dan aturan asosiasi sebagai teori pendukung yang valid dan aplikatif.

Studi Sebelumnya sebagai Fondasi

Kajian terhadap studi sebelumnya atau penelitian terdahulu merupakan bagian integral dari proses ilmiah yang tidak hanya memperkuat landasan teoretis dan metodologis suatu penelitian, tetapi juga mengidentifikasi gap atau kekosongan riset yang masih bisa diisi oleh penelitian terkini. Dalam konteks penelitian ini yang berfokus pada penerapan algoritma Apriori untuk menemukan pola pembelian dalam bisnis alat tulis kantor di Kwek Photo, studi sebelumnya memberikan pijakan penting untuk menyusun kerangka berpikir, menetapkan parameter analisis, serta memanfaatkan hasil-hasil yang telah ada sebagai pembanding serta sumber validasi.

Dalam konteks penelitian yang sedang dilakukan, kajian studi sebelumnya juga menunjukkan kebutuhan akan penelitian yang lebih spesifik ke sektor alat tulis kantor dengan parameter dan model aturan asosiasi yang dioptimalkan untuk konteks tersebut. Gap ini tampak dari keterbatasan riset terdahulu yang sebagian besar fokus pada produk yang berbeda atau belum mengkaji secara mendalam implementasi algoritma Apriori untuk pola pembelian di sektor alat tulis kantor. Oleh karena itu, penelitian Anda mengisi kekosongan tersebut dengan pendekatan

yang lebih terfokus dan aplikasi parameter support dan confidence yang disesuaikan dengan karakteristik data penjualan di Kwek Photo.

Pembandingan dan analisis kritis terhadap berbagai studi yang sudah ada juga membantu menetapkan validitas metode yang dipilih serta memberikan wawasan mengenai potensi kelemahan dan tantangan yang akan dihadapi. Misalnya, beberapa studi mengindikasikan bahwa meskipun algoritma Apriori cepat dalam menemukan pola, pengaturan parameter yang kurang tepat dapat mempengaruhi hasil, sehingga perlu dilakukan perancangan eksperimen parameter yang seksama.

Lebih jauh lagi, kajian terhadap literatur juga memperlihatkan bahwa integrasi antara penemuan pola asosiasi dengan strategi bisnis nyata seperti pengaturan tata letak produk, promosi silang, dan pengembangan sistem rekomendasi produk adalah langkah penting yang sering kurang dioptimalkan dalam studi terdahulu. Penelitian Anda kemudian dapat memberikan kontribusi yang lebih aplikatif dengan menguji implementasi hasil analisis algoritma Apriori ke dalam tindakan bisnis konkret di lapangan.

Dengan demikian, studi sebelumnya tidak saja menjadi landasan teori dan metodologi, tetapi juga berfungsi sebagai pembanding kritis yang memungkinkan Anda menempatkan penelitian ini dalam konteks pengembangan ilmu serta memenuhi kebaruan riset. Ini memberikan kekuatan argumentasi dan memperkuat posisi ilmiah hasil penelitian yang nanti akan dipaparkan dalam bagian hasil dan pembahasan.

BAB 3

Metode Penelitian dan Perancangan

Jenis Data dan Sumbernya

Dalam konteks penelitian ini, persiapan dan pembersihan data sebagai bagian dari tahapan preprocessing merupakan fondasi utama yang harus dilakukan secara sistematis agar proses analisis data menggunakan algoritma Apriori dapat memberikan hasil yang valid dan bermakna. Proses preprocessing berfokus pada transformasi data transaksi mentah menjadi dataset yang bersih, konsisten, dan siap pakai, sehingga memudahkan algoritma dalam menemukan pola asosiasi yang akurat.

Tahap pertama dalam preprocessing adalah pembersihan data (data cleaning). Data mentah yang diperoleh dari sistem penjualan biasanya memiliki berbagai permasalahan seperti data hilang, duplikasi, dan inkonsistensi, yang jika tidak ditangani dapat mengakibatkan bias dan kesalahan interpretasi hasil analisis. Oleh karena itu, langkah awal adalah melakukan identifikasi dan penanganan data hilang yang mungkin muncul, misalnya transaksi yang tidak lengkap yang tidak mencantumkan



item atau jumlah pembelian. Pendekatannya dapat berupa penghapusan data jika jumlahnya sedikit agar tidak mempengaruhi keseluruhan analisis, atau pengisian nilai yang hilang menggunakan metode imputasi yang relevan seperti pengisian dengan modus dari atribut yang bersangkutan.

Selanjutnya adalah penghapusan data duplikat, khususnya transaksi yang tercatat lebih dari sekali. Kehadiran duplikasi ini akan menimbulkan distorsi frekuensi item yang berujung pada pola asosiasi yang tidak akurat. Oleh sebab itu, seluruh transaksi diidentifikasi untuk memastikan tidak ada pengulangan data. Selain itu, koreksi inkonsistensi data kategori juga menjadi perhatian penting. Contohnya adalah ketidakseragaman penulisan nama produk, misalnya mengapa ada variasi “Pulpen” dan “pulpen” yang merujuk pada produk yang sama. Standarisasi nama produk sangat penting agar algoritma dapat mengidentifikasi setiap item secara konsisten tanpa adanya fragmentasi akibat perbedaan penulisan.

Pada tahap ini, validasi nilai numerik juga dilakukan untuk memastikan data seperti harga dan jumlah penjualan berada dalam rentang yang logis. Transaksi yang mengandung nilai negatif atau nol, yang tidak relevan dalam konteks penjualan, harus dikoreksi atau dihapus agar tidak mempengaruhi analisis lebih lanjut.

Setelah data berhasil dibersihkan, tahap berikutnya adalah proses transformasi data. Transformasi ini meliputi konversi data dari format tabular ke format yang lebih sesuai dengan kebutuhan algoritma Apriori, yaitu format transaksi berbentuk kumpulan itemset. Data yang awalnya terdiri dari beberapa kolom seperti kode transaksi, produk, jumlah, dan harga, harus diubah menjadi format dimana tiap transaksi berisi daftar item yang dibeli secara bersamaan. Format ini memudahkan algoritma dalam memproses frequent itemset dan aturan asosiasi.

Selain itu, langkah pengkodean item menjadi bagian dari transformasi ini. Nama produk dikonversi menjadi kode numerik atau string yang konsisten guna meningkatkan efisiensi pemrosesan dan mengurangi potensi error dalam identifikasi item. Jika diperlukan, produk-produk yang memiliki variasi kecil atau jenis serupa dapat digabungkan dalam satu

kategori agar analisis tidak terlalu fragmented, misalnya mengelompokkan semua jenis pulpen dalam satu kategori “Pulpen” tanpa membedakan warna atau merek spesifiknya.

Meski algoritma Apriori tidak memerlukan normalisasi numerik seperti metode lain, penting untuk memastikan bahwa seluruh data kategorikal sudah dalam bentuk seragam dan siap diproses sesuai yang dibutuhkan.

Langkah terakhir dalam persiapan adalah memastikan dataset final yang akan digunakan benar-benar bebas dari noise dan error. Dataset ini berupa kumpulan transaksi yang masing-masing berisi itemset—daftar produk yang dibeli bersama dalam satu transaksi. Kualitas dataset ini akan sangat mempengaruhi performa algoritma dalam menemukan pola yang relevan dan valid. Pada tahap ini, verifikasi ulang dilakukan untuk memastikan tidak ada transaksi kosong, item tidak valid, atau kesalahan data lain yang terlewatkan

Selain itu, meskipun bukan bagian dari preprocessing, persiapan parameter algoritma seperti nilai minimum support dan confidence sangat dipengaruhi oleh kualitas data yang telah dibersihkan dan dipersiapkan. Parameter ini akan menentukan sensitivitas aturan asosiasi yang ditemukan.

Proses persiapan dan pembersihan data tersebut sangat menentukan keberhasilan penelitian, khususnya dalam konteks menemukan pola pembelian yang bermakna dan dapat dijadikan sebagai dasar pengambilan keputusan bisnis di Kwek Photo. Tanpa proses yang teliti ini, risiko mendapatkan pola palsu dan misleading hasil analisis akan sangat tinggi, yang tentu saja akan merugikan strategi bisnis yang diambil berdasarkan hasil tersebut.

Tabel 2 Variabel pada Dataset

Nama Variabel	Deskripsi	Tipe Data
Kode Transaksi	Identifikasi unik setiap transaksi	Numerik
Produk	Nama produk yang dibeli	Kategorikal
Satuan	Unit pengukuran produk (contoh: lembar, pcs)	Kategorikal
Harga	Harga per unit dari produk	Numerik
Jumlah penjualan	Total jumlah unit yang terjual untuk setiap item	Numerik

Tabel 2 menyajikan uraian rinci mengenai variabel-variabel yang terdapat dalam dataset transaksi penjualan yang digunakan sebagai dasar analisis pada penelitian ini. Tabel ini mengelompokkan variabel berdasarkan nama variabel, deskripsi fungsinya, dan tipe data yang mendasari karakteristik data tersebut. Variabel pertama yaitu kode transaksi, yang berperan sebagai identifer unik setiap aksi pembelian dalam basis data dan bertipe data numerik, memungkinkan pemisahan transaksi satu dengan lainnya secara jelas. Variabel produk memberikan informasi spesifik tentang nama barang yang dibeli, dan tersaji sebagai data kategorikal karena berupa nama-nama produk yang berbeda. Selanjutnya, variabel satuan menjelaskan bentuk pengukuran produk, misalnya per lembar, per pack, atau per buah, juga berkategori sebagai data kategorikal.

Variabel harga dalam tabel ini merepresentasikan nilai satuan moneter dari produk terkait, yang bertipe numerik dan memengaruhi analisis nilai transaksi serta potensi keuntungan. Variabel jumlah penjualan menunjukkan kuantitas unit barang yang terjual dalam setiap transaksi, yang juga bertipe numerik, menjadi kunci dalam menghitung frekuensi dan volume penjualan. Melalui tabel ini, pembaca dapat memahami struktur data penelitian yang sangat rinci dan terorganisir dengan baik, menjadi landasan bagi proses data mining menggunakan algoritma Apriori.

Tabel 2 sangat penting karena menjelaskan setiap komponen data secara sistematis yang memudahkan analisis, interpretasi hasil, sekaligus menjadi panduan saat melakukan tahap pengolahan dan validasi data. Dengan memahami variabel ini, pembaca dapat menilai bagaimana data tersebut digunakan untuk mengekstrak pola asosiasi yang relevan dan implementasinya dalam pengambilan keputusan bisnis. Keseluruhan isi tabel memperlihatkan kesiapan data sebagai fondasi kuat untuk penelitian dan kegunaan praktisnya dalam optimasi strategi bisnis Kwek Photo.

Persiapan dan Pembersihan Data

Dalam konteks penelitian ini, persiapan dan pembersihan data sebagai bagian dari tahapan preprocessing merupakan fondasi utama yang harus dilakukan secara sistematis agar proses analisis data menggunakan algoritma Apriori dapat memberikan hasil yang valid dan bermakna. Proses preprocessing berfokus pada transformasi data transaksi mentah menjadi dataset yang bersih, konsisten, dan siap pakai, sehingga memudahkan algoritma dalam menemukan pola asosiasi yang akurat.

Tahap pertama dalam preprocessing adalah pembersihan data (data cleaning). Data mentah yang diperoleh dari sistem penjualan biasanya memiliki berbagai permasalahan seperti data hilang, duplikasi, dan inkonsistensi, yang jika tidak ditangani dapat mengakibatkan bias dan kesalahan interpretasi hasil analisis. Oleh karena itu, langkah awal adalah melakukan identifikasi dan penanganan data hilang yang mungkin muncul, misalnya transaksi yang tidak lengkap yang tidak mencantumkan item atau jumlah pembelian. Pendekatannya dapat berupa penghapusan data jika jumlahnya sedikit agar tidak mempengaruhi keseluruhan analisis, atau pengisian nilai yang hilang menggunakan metode imputasi yang relevan seperti pengisian dengan modus dari atribut yang bersangkutan.

Selanjutnya adalah penghapusan data duplikat, khususnya transaksi yang tercatat lebih dari sekali. Kehadiran duplikasi ini akan menimbulkan distorsi frekuensi item yang berujung pada pola asosiasi yang tidak akurat.

Oleh sebab itu, seluruh transaksi diidentifikasi untuk memastikan tidak ada pengulangan data. Selain itu, koreksi inkonsistensi data kategori juga menjadi perhatian penting. Contohnya adalah ketidakseragaman penulisan nama produk, misalnya mengapa ada variasi “Pulpen” dan “pulpen” yang merujuk pada produk yang sama. Standarisasi nama produk sangat penting agar algoritma dapat mengidentifikasi setiap item secara konsisten tanpa adanya fragmentasi akibat perbedaan penulisan.

Pada tahap ini, validasi nilai numerik juga dilakukan untuk memastikan data seperti harga dan jumlah penjualan berada dalam rentang yang logis. Transaksi yang mengandung nilai negatif atau nol, yang tidak relevan dalam konteks penjualan, harus dikoreksi atau dihapus agar tidak mempengaruhi analisis lebih lanjut.

Tabel 3 Dataset Transaksi Penjualan Alat Tulis Kantor

Kode Transaksi	Produk	Satuan	Harga (Rp)	Jumlah penjualan
1	kertas hvs	lembar	165	15
2	kertas foto	lembar	3.000	11
3	kertas folio	lembar	500	12
4	kertas jeruk	lembar	1.000	15
5	kertas BC	lembar	2.000	9
6	kertas linen BC	lembar	5.000	16
7	kertas CV kosong	lembar	4.000	10
8	kertas NCR	lembar	750	11
9	laminating PVC card	pcs	7.500	5
10	plastik laminating	pcs	6.000	15
11	plastik mika jilid	pcs	2.500	20
12	map plastik tali	pcs	3.000	8

Kode Transaksi	Produk	Satuan	Harga (Rp)	Jumlah penjualan
13	map kancing	pcs	3.500	5
14	map zipper	pcs	4.000	8
15	map bisnis file	pcs	8.000	4
16	map dokumen	pcs	6.500	3
17	map amplop 310	pcs	5.000	10
18	buku ekspedisi	pcs	10.000	7
19	buku folio garda	pcs	7.000	6
20	buku milimeter blok	pcs	5.500	6
21	bon faktur	pcs	2.000	10
22	buku binder	pcs	15.000	10
23	refill binder	pcs	5.000	5
24	binder klip	pcs	2.500	10
25	paper klip	pcs	1.500	20
26	kwitansi	pcs	1.000	10
27	stiker label	pcs	3.000	12
28	amplop	pcs	500	17
29	lem kertas	pcs	2.000	10
30	lem fox	pcs	3.500	11
31	tipe-x	pcs	2.000	9
32	rautan	pcs	1.000	19
33	penghapus pensil	pcs	500	10
34	staples	pcs	2.000	7
35	stamp pad	pcs	3.000	6

Kode Transaksi	Produk	Satuan	Harga (Rp)	Jumlah penjualan
36	tinta stempel	pcs	5.000	3
37	pulpen	pcs	1.000	12
38	spidol	pcs	3.000	8
39	block notes	pcs	4.500	14
40	lakban	pcs	3.500	6
41	kertas karton	lembar	2.000	13
42	kertas manila	lembar	1.500	17

Tabel 3 menyajikan dataset yang digunakan dalam penelitian ini, yang berfokus pada penerapan data mining untuk menganalisis pola pembelian alat tulis kantor di Kwek Photo. Dataset ini terdiri dari 42 entri yang mencakup informasi penting seperti kode transaksi, produk, satuan, dan harga. Kode transaksi berfungsi sebagai identifikasi unik untuk setiap transaksi, sedangkan produk mencakup berbagai jenis barang yang dijual, seperti kertas, alat tulis, dan perlengkapan kantor lainnya. Satuan menunjukkan unit pengukuran yang digunakan untuk setiap produk, baik dalam bentuk pcs maupun lembar. Harga mencerminkan nilai ekonomis dari setiap item, yang penting untuk analisis lebih lanjut. Dengan dataset ini, analisis dapat dilakukan untuk mengidentifikasi pola pembelian dan merumuskan strategi pemasaran yang lebih efektif, sehingga Kwek Photo dapat meningkatkan efisiensi operasional dan daya saing di pasar alat tulis kantor.

Setelah data berhasil dibersihkan, tahap berikutnya adalah proses transformasi data. Transformasi ini meliputi langkah pengkodean item menjadi bagian dari transformasi ini. Nama produk dikonversi menjadi kode numerik atau string yang konsisten guna meningkatkan efisiensi pemrosesan dan mengurangi potensi error dalam identifikasi item. Jika diperlukan, produk-produk yang memiliki variasi kecil atau jenis

serupa dapat digabungkan dalam satu kategori agar analisis tidak terlalu fragmented, misalnya mengelompokkan semua jenis pulpen dalam satu kategori “Pulpen” tanpa membedakan warna atau merek spesifiknya.

Tabel 4 Data ATK dengan Kode

Kode	Nama Alat Tulis Kantor
ATK1	Kertas HVS
ATK2	Kertas Foto
ATK3	Kertas Folio
ATK4	Kertas Jeruk
ATK5	Kertas BC
ATK6	Kertas Linen BC
ATK7	Kertas CV Kosong
ATK8	Kertas NCR
ATK9	Laminating PVC Card
ATK10	Plastik Laminating
ATK11	Plastik Mika Jilid
ATK12	Map Plastik Tali
ATK13	Map Kancing
ATK14	Map Zipper
ATK15	Map Bisnis File
ATK16	Map Dokumen
ATK17	Map Amplop 310
ATK18	Buku Expedisi
ATK19	Buku Folio Garda
ATK20	Buku Milimeter Blok
ATK21	Bon Faktur
ATK22	Buku Binder
ATK23	Refill Binder
ATK24	Binder Klip
ATK25	Paper Klip
ATK26	Kwitansi
ATK27	Stiker Label
ATK28	Amplop

Kode	Nama Alat Tulis Kantor
ATK29	Lem Kertas
ATK30	Lem Fox
ATK31	Tipe-X
ATK32	Rautan
ATK33	Penghapus Pensil
ATK34	Staples
ATK35	Stamp Pad
ATK36	Tinta Stempel
ATK37	Pulpen
ATK38	Spidol
ATK39	Block Notes
ATK40	Lakban
ATK41	Kertas Karton
ATK42	Kertas Manila

Selain itu melakukan transformasi data dalam format tabular lebih sesuai dengan kebutuhan algoritma Apriori, yaitu format transaksi berbentuk kumpulan itemset. Data yang awalnya terdiri dari beberapa kolom seperti kode transaksi, produk, jumlah, dan harga, harus diubah menjadi format dimana tiap transaksi berisi daftar item yang dibeli secara bersamaan. Format ini memudahkan algoritma dalam memproses frequent itemset dan aturan asosiasi.

Tabel 5 Tabular

Tanggal	ATK1	ATK2	ATK41	ATK42
01 Januari 2025	1	0	1	1
02 Januari 2025	0	1	1	0
03 Januari 2025	1	1	1	0
04 Januari 2025	0	0	1	1
05 Januari 2025	0	1	0	1
06 Januari 2025	0	0	1	0
07 Januari 2025	1	1	0	1
08 Januari 2025	0	1	0	0

Tanggal	ATK1	ATK2	ATK41	ATK42
09 Januari 2025	0	0	1	0
10 Januari 2025	0	1	0	1
11 Januari 2025	1	0	0	0
12 Januari 2025	0	0	1	0
13 Januari 2025	0	1	0	1
14 Januari 2025	0	0	0	0
15 Januari 2025	1	0	0	1
16 Januari 2025	1	1	1	1
17 Januari 2025	1	0	0	1
18 Januari 2025	1	1	1	1
19 Januari 2025	0	0	0	1
20 Januari 2025	0	0	0	0
21 Januari 2025	1	1	1	0
22 Januari 2025	1	0	0	1
23 Januari 2025	0	0	0	0
24 Januari 2025	0	0	0	1
25 Januari 2025	1	0	1	1
26 Januari 2025	0	1	0	0
27 Januari 2025	1	0	0	1
28 Januari 2025	1	0	1	0
29 Januari 2025	0	0	0	1
30 Januari 2025	1	0	1	0
31 Januari 2025	1	0	0	1

Meski algoritma Apriori tidak memerlukan normalisasi numerik seperti metode lain, penting untuk memastikan bahwa seluruh data kategorikal sudah dalam bentuk seragam dan siap diproses sesuai yang dibutuhkan.

Langkah terakhir dalam persiapan adalah memastikan dataset final yang akan digunakan benar-benar bebas dari noise dan error. Dataset ini berupa kumpulan transaksi yang masing-masing berisi itemset daftar produk yang dibeli bersama dalam satu transaksi. Kualitas dataset ini akan sangat mempengaruhi performa algoritma dalam menemukan pola

yang relevan dan valid. Pada tahap ini, verifikasi ulang dilakukan untuk memastikan tidak ada transaksi kosong, item tidak valid, atau kesalahan data lain yang terlewatkan.

Selain itu, meskipun bukan bagian dari preprocessing, persiapan parameter algoritma seperti nilai minimum support dan confidence sangat dipengaruhi oleh kualitas data yang telah dibersihkan dan dipersiapkan. Parameter ini akan menentukan sensitivitas aturan asosiasi yang ditemukan.

Proses persiapan dan pembersihan data tersebut sangat menentukan keberhasilan penelitian, khususnya dalam konteks menemukan pola pembelian yang bermakna dan dapat dijadikan sebagai dasar pengambilan keputusan bisnis di Kwek Photo. Tanpa proses yang teliti ini, risiko mendapatkan pola palsu dan misleading hasil analisis akan sangat tinggi, yang tentu saja akan merugikan strategi bisnis yang diambil berdasarkan hasil tersebut.

Merancang Strategi Analisis Data

Strategi analisis data merupakan aspek kunci dalam penelitian yang bertujuan untuk menggali pola-pola signifikan pada data transaksi menggunakan algoritma Apriori. Penyusunan strategi ini harus dilakukan secara sistematis dan logis, dimulai dari pemahaman mendalam terhadap karakteristik data hingga penerapan metodologi yang tepat guna menghasilkan alat analisis yang efektif dan akurat.

Langkah pertama dalam merancang strategi adalah mengidentifikasi tujuan utama analisis, yaitu menemukan asosiasi atau pola keterkaitan antar item dalam data penjualan produk di Kwek Photo. Tujuan ini mengarahkan seluruh proses desain penelitian, khususnya dalam memilih dan menyusun teknik serta parameter yang akan digunakan dalam algoritma Apriori sebagai inti dari metodologi analisis.

Setelah tujuan ditetapkan, langkah berikutnya adalah menyiapkan dataset yang telah melalui tahapan preprocessing seperti pembersihan

dan transformasi data agar siap untuk dianalisis. Dataset ini kemudian dianalisis dengan algoritma Apriori yang memerlukan struktur data dalam bentuk transaksi itemset. Oleh karena itu, strategi analisis ini menyusun langkah logis yang mencakup parameter penting algoritma, yaitu minimum support dan minimum confidence, yang akan menentukan kualitas dan relevansi aturan asosiasi yang dihasilkan. Nilai parameter ini harus dipilih berdasarkan pertimbangan empiris dan domain knowledge, agar menghasilkan pola yang meaningful dan tidak terlalu banyak aturan yang tidak relevan atau noise.

Selain itu, metodologi analisis harus mencakup langkah verifikasi dan validasi hasil yang diperoleh dari algoritma. Hal ini mencakup evaluasi kekuatan aturan asosiasi melalui ukuran support, confidence, dan lift, serta interpretasi hasil berdasarkan konteks bisnis yang menjalankan penelitian. Proses interpretasi ini tidak hanya bersifat teknis tetapi juga membutuhkan pemahaman bisnis agar hasil analisis dapat digunakan secara efektif dalam pengambilan keputusan.

Strategi analisis data juga harus mengakomodasi mekanisme iteratif. Misalnya, bila hasil awal menunjukkan terlalu banyak aturan dengan confidence rendah atau aturan yang tidak praktis, maka nilai support dan confidence harus disesuaikan kembali dengan proses tuning parameter yang sistematis hingga diperoleh aturan yang paling optimal. Metode tuning ini bisa menjadi langkah esensial dalam strategi agar hasil analisis benar-benar berguna dan usable dalam skenario nyata.

Selanjutnya, tahapan visualisasi hasil analisis juga menjadi bagian dari strategi ini. Visualisasi data dan aturan asosiasi dapat digunakan untuk memperjelas pola yang ditemukan, misalnya dengan menggunakan diagram asosiasi atau grafik frekuensi itemset sehingga pembaca maupun pemangku kepentingan bisnis dapat dengan mudah memahami informasi yang dihasilkan. Pendekatan visual ini juga mendukung proses komunikasi hasil riset secara efektif.

Strategi ini juga menyertakan pertimbangan teknis terkait implementasi algoritma Apriori, seperti pemilihan perangkat lunak

analisis yang sesuai, penanganan data besar (big data) jika diperlukan, serta optimasi runtime algoritma agar proses analisis dapat berjalan efisien dan dalam waktu yang terjangkau.

Selain langkah teknis dan metodologis, perancangan strategi analisis data harus menjelaskan prosedur dokumentasi dan pelaporan hasil analisis secara sistematis dan terstruktur. Ini termasuk penjelasan mengenai bagaimana dokumentasi aturan asosiasi yang ditemukan akan dilakukan, serta bagaimana interpretasi hasil akan diintegrasikan ke dalam laporan penelitian sehingga memperkuat argumentasi dan rekomendasi yang diajukan.

Dengan menyusun strategi analisis data secara runut dan sistematis, penelitian ini tidak hanya mampu menghasilkan pola-pola asosiasi yang valid tetapi juga menyediakan metodologi yang dapat direplikasi maupun dikembangkan pada penelitian selanjutnya, baik dengan algoritma Apriori maupun algoritma asosiasi lainnya.

Menentukan Parameter Kunci Analisis

Dalam penelitian data mining yang menggunakan algoritma Apriori, penentuan parameter kunci seperti nilai minimum support dan minimum confidence sangat krusial untuk menghasilkan aturan asosiasi yang valid, bermakna, dan aplikatif dalam konteks bisnis. Support dan confidence adalah dua metrik utama yang menjadi fondasi dalam proses penggalian pola dari data transaksi. Pemilihan nilai threshold untuk kedua parameter ini harus didasari oleh pertimbangan yang matang baik dari aspek teknis algoritma maupun relevansi bisnis.

Minimum support adalah ambang batas yang menentukan seberapa sering sebuah itemset muncul dalam data transaksi agar dianggap signifikan dan layak diproses lebih lanjut. Support diukur sebagai persentase frekuensi kemunculan itemset tersebut dibandingkan dengan total transaksi. Dalam konteks penelitian ini, ditetapkan support minimum sebesar 17%. Nilai ini menandakan bahwa itemset harus muncul sedikitnya dalam 17% dari

seluruh transaksi agar dapat dianggap memenuhi syarat untuk dianalisis sebagai pola yang potensial. Pemilihan angka 17% merupakan kompromi antara keluwesan menemukan aturan baru dan penghindaran dari aturan yang terlalu jarang sehingga menjadi noise atau kurang relevan.

Nilai support yang terlalu rendah bisa membuat algoritma memproses banyak aturan yang sifatnya minor dan kurang menguntungkan secara bisnis, sehingga membebani proses komputasi dan menghasilkan output yang tidak fokus. Sebaliknya, support yang terlalu tinggi berisiko mengabaikan aturan potensial yang bisa memberikan insight baru. Dengan nilai support 17%, algoritma dapat mengeliminasi aturan dengan frekuensi kemunculan yang sangat rendah atau outlier, sehingga fokus analisis tetap pada itemset populer yang memberi nilai strategis bagi bisnis.

Minimum confidence mengukur tingkat kepercayaan atau probabilitas bahwa item Y akan dibeli apabila item X sudah dibeli (dalam aturan $X \rightarrow Y$). Confidence adalah indikator kekuatan hubungan antar item. Nilai confidence minimum yang dipilih adalah 50%, yang berarti aturan harus memiliki tingkat kepercayaan setidaknya separuh kali transaksi dimana X muncul, untuk dianggap signifikan. Dengan threshold ini, hanya aturan dengan korelasi cukup kuat yang diterima sebagai hasil. Threshold 50% ini pragmatis dan sering digunakan dalam konteks bisnis untuk menjaga keseimbangan antara jumlah aturan yang dihasilkan dan kualitasnya.

Confidence yang terlalu rendah menghasilkan banyak aturan lemah yang sulit diinterpretasi dan tidak handal dalam pengambilan keputusan. Confidence tinggi akan mempersempit aturan sehingga hanya yang benar-benar kuat yang dianggap valid. Dengan adanya batasan ini, fokus penelitian dapat tertuju pada aturan-aturan yang secara statistik dan praktis mampu mendukung strategi bisnis.

Penentuan parameter ini sejalan dengan tujuan penelitian untuk menemukan pola yang tidak hanya secara statistik signifikan, tetapi juga secara bisnis dapat diterapkan untuk meningkatkan pendapatan, loyalitas pelanggan, dan efisiensi penataan produk.

BAB 4



Implementasi dan Pembahasan

Mengaplikasikan Algoritma Apriori Secara Manual

Pada tahap implementasi penelitian data mining menggunakan algoritma Apriori, pengaplikasian secara manual terhadap dataset merupakan proses krusial agar gambaran proses algoritma dapat dipahami secara konseptual dan diterapkan dengan tepat dalam situasi nyata. Algoritma Apriori digunakan untuk menemukan pola atau asosiasi yang sering muncul dalam data transaksi, yakni hubungan antara satu set item yang secara rutin muncul bersama dalam transaksi pelanggan.

Langkah pertama dalam penerapan algoritma Apriori secara manual adalah penentuan dataset transaksi yang akan dianalisis. Dataset tersebut harus berupa kumpulan data transaksi yang lengkap dan sudah melalui proses pembersihan serta transformasi agar formatnya sesuai dengan kebutuhan algoritma, yakni dalam bentuk daftar item yang dibeli berdasarkan setiap transaksi. Setiap transaksi dianggap sebagai satu himpunan item.

Selanjutnya adalah menentukan nilai minimum support dan minimum confidence yang menjadi parameter utama dalam penggalian aturan

asosiasi. Support mengacu pada frekuensi relatif kemunculan itemset dalam seluruh transaksi, sementara confidence mengukur kekuatan hubungan antar item pada aturan asosiasi yang dihasilkan. Nilai minimum support dan confidence ini menjadi acuan untuk mensaring itemset dan aturan yang dianggap signifikan dan relevan, sehingga data yang diproses tidak berlebihan dan tetap fokus kepada pola yang bermakna.

Perhitungan Manual Support

Support adalah salah satu metrik utama dalam algoritma Apriori yang digunakan untuk mengukur seberapa sering suatu item atau kombinasi item muncul dalam transaksi. Nilai support memberikan wawasan tentang popularitas suatu produk dalam keseluruhan transaksi, sehingga membantu untuk mengidentifikasi pola pembelian konsumen. Rumus perhitungan support adalah sebagai berikut:

$$Support(X) = \frac{\text{Jumlah transaksi yang mengandung } X}{\text{Total transaksi}}$$

Dimana:

- Support(A): Merupakan nilai yang menggambarkan seberapa sering item A muncul di dalam keseluruhan kumpulan data transaksi.
- Jumlah transaksi yang mengandung item A: Menunjukkan banyaknya transaksi dalam dataset yang memuat item A.
- Total jumlah transaksi: Adalah keseluruhan transaksi yang tercatat dalam dataset.

Diketahui jumlah transaksi sebanyak 31 transaksi dan produk ATK sebanyak 42 data (tabel 5 tabular). Nilai support dinyatakan dalam bentuk persentase dan menunjukkan proporsi transaksi yang mengandung item tertentu. Perhitungan dimulai dari masing-masing produk ATK (1 itemset):

1. ATK1 = Kertas HVS

Banyak transaksi yang melibatkan kertas HVS = 15.

48%

- 2. ATK2 = Kertas Foto
Banyak transaksi yang melibatkan kertas HVS = 11.
35%
- 3. ATK3 = Kertas Folio
Banyak transaksi yang melibatkan kertas HVS = 12.
39%
- 4. ATK4 = Kertas Jeruk
Banyak transaksi yang melibatkan kertas HVS = 15.
48%
- 5. ATK5 = Kertas BC
Banyak transaksi yang melibatkan kertas HVS = 9.
29%

Setelah selesai dilakukan perhitungan keseluruhan produk, selanjutnya memilih item mana yang memenuhi kriteria. Berikut adalah hasil dari perhitungan keseluruhannya

Tabel 6. Hasil perhitungan support

Produk	Support
kertas hvs	48%
kertas foto	35%
kertas folio	39%
kertas jeruk	48%
kertas BC	29%
kertas linen BC	52%
kertas CV kosong	32%
kertas NCR	35%
plastik laminating	48%
plastik mika jilid	65%
map plastik tali	26%
map zipper	26%
map amplop 310	32%

Produk	Support
buku ekspedisi	23%
buku folio garda	19%
buku milimeter blok	19%
bon faktur	32%
buku binder	32%
binder klip	32%
paper klip	65%
kwitansi	32%
stiker label	39%
amplop	52%
lem kertas	32%
lem fox	35%
tipe-x	29%
rautan	61%
penghapus pensil	32%
staples	23%
stamp pad	19%
pulpen	39%
spidol	26%
block notes	45%
lakban	19%
kertas karton	42%
kertas manila	55%

Dalam hal ini, nilai minimum support ditentukan sebesar 17%. Oleh karena itu, anggota himpunan 1-itemset yang memiliki nilai support kurang dari 17% akan dikeluarkan dari proses iterasi berikutnya. Sedangkan anggota himpunan dengan nilai support sama dengan atau lebih besar dari 17% akan masuk ke dalam large 1-itemset yang memenuhi syarat minimum support.

Perhitungan manual dengan kombinasi 2 produk (2 itemset):

1. ATK1 = Kertas HVS dan ATK2 = Kertas Foto.

Banyak transaksi yang melibatkan ATK1 dan ATK2 = 5.

16%

2. ATK1 = Kertas HVS dan ATK3 = Kertas Folio.
Banyak transaksi yang melibatkan ATK1 dan ATK3 = 7.

23%

3. ATK1 = Kertas HVS dan ATK4 = Kertas Jeruk.
Banyak transaksi yang melibatkan ATK1 dan ATK4 = 6.

19%

4. ATK1 = Kertas HVS dan ATK5 = Kertas BC.
Banyak transaksi yang melibatkan ATK1 dan ATK5 = 4.

13%

5. ATK1 = Kertas HVS dan ATK6 = Kertas Linen BC.
Banyak transaksi yang melibatkan ATK1 dan ATK6 = 8.

26%

Setelah selesai dilakukan perhitungan keseluruhan produk, selanjutnya memilih item mana yang memenuhi kriteria. Berikut adalah hasil dari perhitungan keseluruhannya:

Tabel 6. Hasil perhitungan support kombinasi 2 itemset

Kombinasi 2 itemset	Support
kertas hvs-kertas folio	23%
kertas hvs-kertas jeruk	19%
kertas hvs-kertas linen bc	26%
kertas hvs-plastik laminating	23%
kertas hvs-plastik mika jilid	35%
.....
.....
pulpen-block notes	26%
pulpen-kertas manila	26%
lakban-kertas manila	19%
block notes-kertas manila	26%

Kombinasi 2 itemset	Support
block notes-kertas karton	23%

Perhitungan manual dengan kombinasi 3 produk yang dibeli secara bersamaan (3 itemset):

1. ATK1 = Kertas HVS, ATK2 = Kertas Foto dan ATK3 = Kertas Folio.
Banyak transaksi yang melibatkan ATK1, ATK2 dan ATK3 = 7.
23%
2. ATK1 = Kertas HVS, ATK2 = Kertas Foto dan ATK4 = Kertas Jeruk.
Banyak transaksi yang melibatkan ATK1, ATK2 dan ATK4 = 10.
32%
3. ATK1 = Kertas HVS, ATK2 = Kertas Foto dan ATK5 = Kertas BC.
Banyak transaksi yang melibatkan ATK1, ATK2 dan ATK5 = 5.
16%
4. ATK1 = Kertas HVS, ATK2 = Kertas Foto dan ATK6 = Kertas BC.
Banyak transaksi yang melibatkan ATK1, ATK2 dan ATK6 = 10.
32%
5. ATK1 = Kertas HVS, ATK2 = Kertas Foto dan ATK7 = Kertas Folio.
Banyak transaksi yang melibatkan ATK1, ATK2 dan ATK7 = 10.
32%

Perhitungan manual dilakukan hanya sampai 5 sampel perhitungan. Untuk perhitungan selanjutnya dilakukan dengan cara yang sama dengan menyesuaikan transaksi dan produknya. Berikut adalah hasil dari perhitungan kombinasi 3 itemset:

Tabel 7. Hasil perhitungan support kombinasi 3 itemset

Kombinasi 3 itemset	Support
kertas hvs-kertas foto-kertas folio	23%
kertas hvs-kertas foto-kertas jeruk	32%
kertas hvs-kertas foto-kertas linen BC	32%
kertas hvs-kertas foto-kertas CV kosong	32%

Kombinasi 3 itemset	Support
kertas hvs-kertas foto-kertas NCR	23%
.....
.....
kertas hvs-kertas linen BC-spivol	32%
kertas hvs-kertas linen BC-block notes	35%
kertas hvs-kertas linen BC-lakban	29%
kertas hvs-kertas linen BC-kertas karton	29%
kertas hvs-kertas linen BC-kertas manila	35%

Maksud dari kombinasi pertama (kertas hvs-kertas foto-kertas folio) yaitu jika membeli kertas hvs dan kertas foto maka akan membeli kertas folio. Begitu seterusnya dalam pembacaan kombinasinya.

Perhitungan Manual Confidence

Confidence adalah metrik utama dalam algoritma Apriori yang digunakan untuk mengevaluasi kemungkinan suatu item Y dibeli ketika item X sudah dibeli. Metrik ini digunakan untuk membangun aturan asosiasi yang dapat membantu dalam memahami pola pembelian pelanggan dan mendukung strategi pemasaran berbasis data. Rumus perhitungan confidence adalah sebagai berikut:

Dimana:

$$Confidence(X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung } (X \cup Y)}{\text{Jumlah transaksi yang mengandung } X} \times 100\%$$

- X merupakan item awal atau antecedent dalam aturan asosiasi.
- Y adalah item tujuan atau consequent yang mengikuti dari X.
- $X \cap Y$ menunjuk pada transaksi-transaksi yang memuat kedua item X dan Y secara bersamaan.
- Support(X) menandakan proporsi transaksi yang hanya memuat item X. Nilai minimal support yang diterapkan adalah 0,17 atau 17%.

Pada kombinasi 2 item terdapat kombinasi kertas hvs-plastik mika jilid, maka untuk menghitung nilai confidence nya yaitu:

- Total transaksi kertas hvs (x)= 15
- Total transaksi plastik mika jilid (y) = 20
- Total transaksi memuat kertas hvs-plastik mika jilid = 11

$$confidence = \frac{11}{15} \times 100\% = 73\%$$

Pada kombinasi 2 item terdapat kombinasi kertas hvs-plastik mika jilid, maka untuk menghitung nilai confidence nya yaitu:

- Total transaksi plastik mika jilid (x) = 20
- Total transaksi kertas hvs (y)= 15
- Total transaksi memuat kertas hvs-plastik mika jilid = 11

$$confidence = \frac{11}{20} \times 100\% = 55\%$$

Dapat dilihat hasil dari analisis pada item kertas hvs-plastik mika jilid, memiliki nilai support 35% dan confidence 73%. Dimana masing masing nilai melewati batas minimum support (17%) dan confidence (50%) yang artinya jika pelanggan membeli kertas hvs maka 73% kemungkinan pelanggan juga membeli plastic mika jilid.

Pemanfaatan Software RapidMiner dalam Analisis

RapidMiner merupakan salah satu software analitik data yang sangat populer dan banyak digunakan dalam berbagai bidang penelitian serta aplikasi bisnis, khususnya dalam proses penambangan data (data mining). Pemanfaatan RapidMiner dalam analisis data menawarkan pendekatan praktis yang bertujuan mempercepat dan mempermudah proses eksplorasi, pengolahan, serta ekstraksi informasi berharga dari data yang tersedia. Keunggulan utama RapidMiner terletak pada kemampuannya menyediakan antarmuka visual berbasis drag-and-drop yang intuitif, memungkinkan pengguna, baik yang berpengalaman maupun pemula, untuk mengembangkan model analisis data tanpa memerlukan pemrograman rumit.

Dalam konteks penelitian yang melibatkan pengolahan data transaksi atau data besar, seperti studi yang menggunakan algoritma Apriori untuk menemukan pola asosiasi, RapidMiner menjadi alat yang sangat mendukung. Proses analisis yang meliputi pembersihan (data preprocessing), transformasi data, penerapan algoritma, hingga evaluasi hasil dapat dilakukan secara terintegrasi dalam satu platform. Hal ini mengurangi ketergantungan pada berbagai perangkat lunak berbeda, sehingga proses menjadi lebih efisien dan terstruktur. Misalnya, fitur preprocessing di RapidMiner membantu pengguna melakukan seleksi atribut, menghapus data duplikat, mengisi nilai kosong, dan melakukan normalisasi data, yang merupakan langkah awal penting dalam menjamin kualitas analisis.

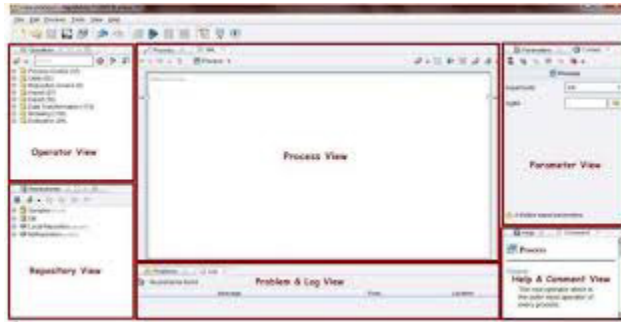
Selanjutnya, RapidMiner menyediakan pustaka algoritma analisis data yang lengkap dan terus diperbarui, termasuk algoritma Apriori yang terkenal untuk penambangan pola asosiasi. Dengan menggunakan RapidMiner, pengguna dapat dengan mudah mengkonfigurasi parameter seperti nilai minimum support dan confidence, yang berpengaruh pada keketatan pola yang disaring dalam proses mining. Kepraktisan ini memungkinkan proses eksplorasi pola pembelian atau hubungan antar item dalam dataset dilakukan tanpa kompleksitas teknis yang tinggi. Output dari algoritma pun dapat divisualisasikan dalam bentuk grafik, diagram, atau tabel yang memperjelas interpretasi hasil.



Gambar 1 Software RapidMiner

Lebih jauh, RapidMiner dilengkapi dengan fitur evaluasi model yang memungkinkan pengguna menilai performa dan keakuratan model analisis. Hal ini penting agar hasil yang diperoleh dapat dipercaya dan relevan dengan tujuan analisis, terutama ketika hasil digunakan untuk pengambilan keputusan bisnis atau rekomendasi strategis. Integrasi

visualisasi yang intuitif juga menjadi aspek kunci yang membantu pemahaman pola dan interpretasi data secara mendalam. Dengan demikian, pengguna dapat melihat tren, cluster, ataupun asosiasi data secara langsung, yang lebih sulit dilakukan bila analisis hanya berbasis teks.



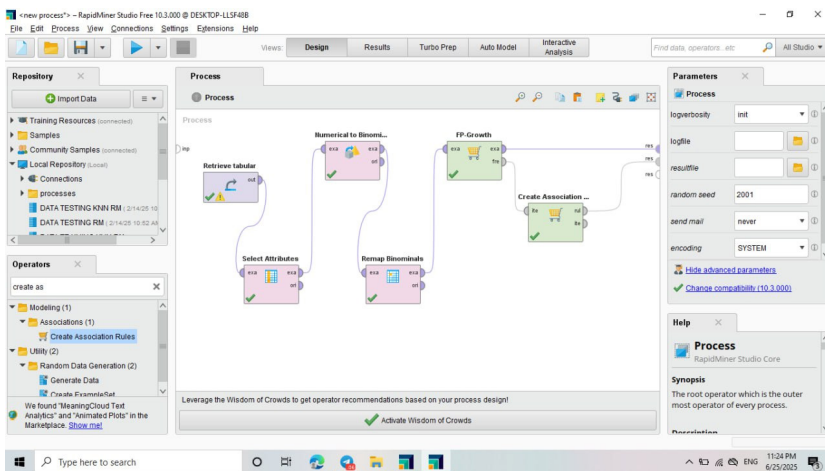
Gambar 2 Menu Software RapidMiner

Kemudahan penggunaan RapidMiner juga membantu mendukung kolaborasi tim penelitian atau tim bisnis karena model dan proses analisis dapat didokumentasikan dan dibagikan dengan mudah. Hal ini menjamin transparansi, konsistensi, dan reproduisibilitas hasil, yang menjadi aspek penting dalam riset akademik maupun lingkungan industri profesional. Selain itu, RapidMiner menyediakan fitur ekspor dan integrasi ke banyak format data dan sistem lain seperti Excel, CSV, dan database, sehingga mempermudah integrasi hasil analisis ke dalam workflow yang lebih luas.

Secara keseluruhan, pemanfaatan software RapidMiner dalam analisis data tidak hanya mempercepat proses analisis tetapi juga meningkatkan kualitas dan kedalaman wawasan yang diperoleh. Dalam konteks penelitian yang memanfaatkan algoritma Apriori, RapidMiner mendorong pendekatan yang lebih sistematis, terstruktur, dan berbasis bukti. Ini sejalan dengan kebutuhan penelitian modern yang menuntut kecepatan dan akurasi dalam mengolah data kompleks serta menghasilkan rekomendasi yang dapat diandalkan.

Langkah Teknis di RapidMiner

Panduan teknis menjalankan algoritma Apriori pada platform RapidMiner dimulai dengan pemahaman alur kerja atau workflow yang mendasari keseluruhan proses analisis data. Gambar yang dilampirkan menampilkan beberapa operator penting dalam RapidMiner yang disusun secara berurutan, mulai dari tahap pengambilan data hingga ke proses pembentukan aturan asosiasi menggunakan algoritma Apriori dengan pendekatan FP-Growth. Setiap langkah dalam workflow ini memiliki fungsi dan parameter yang krusial untuk memastikan keluaran yang akurat dan relevan.



Gambar 3 Parameter dan atribut yang digunakan

Proses analisis data menggunakan RapidMiner, dimulai dari tahap pengambilan data (Retrieve Tabular), seleksi atribut (Select Attributes), transformasi data numerik ke binomial (Numerical to Binomial dan Remap Binominals), hingga penerapan algoritma FP-Growth dan pembentukan aturan asosiasi (Create Association Rules). Visualisasi ini penting untuk membantu pemahaman pengguna tentang proses terstruktur yang diawali dengan persiapan data yang selektif dan transformasi format agar sesuai kebutuhan algoritma.

Setiap blok mewakili operator yang memiliki fungsi spesifik dan rentang pengaturan parameter yang harus diperhatikan agar analisis berjalan optimal. Alur data yang mengalir dari kiri ke kanan menunjukkan proses bertahap dari data mentah menjadi informasi bermakna. Keterangan visual tanda cek hijau menunjukkan setiap langkah berhasil dijalankan tanpa error, menandakan workflow sudah optimal secara teknis.

Langkah pertama pada alur kerja ini adalah operator Retrieve Tabular, yang bertugas mengambil dataset dalam format tabel yang menjadi sumber data untuk analisis. Pada tahap ini, pengguna umumnya memilih dataset yang sudah disiapkan sebelumnya, yang berisi transaksi atau data atribut yang akan diproses. Keberhasilan dan kualitas data yang diambil sangat menentukan hasil akhir analisis, sehingga penting untuk memastikan bahwa data telah bersih dan sesuai dengan kebutuhan.

Setelah pengambilan data, langkah selanjutnya adalah penggunaan operator Select Attributes. Operator ini sangat vital karena berfungsi untuk memilih atribut-atribut atau kolom data yang relevan dan diperlukan dalam proses analisis. Dalam konteks algoritma Apriori, biasanya hanya atribut yang merepresentasikan item transaksi yang dipilih, menghilangkan kolom yang tidak terkait agar tidak mengganggu hasil analisis pola asosiasi. Pengaturan parameter di operator ini memungkinkan pengguna untuk secara manual menentukan kolom yang diikutsertakan atau dikeluarkan dari proses berikutnya, memberikan fleksibilitas dan kontrol yang lebih baik terhadap data input.

Selanjutnya, operator Numerical to Binominal dan Remap Binominals digunakan untuk mengkonversi data numerik menjadi format binomial (biner). Transformasi ini sangat penting karena algoritma Apriori dan FP-Growth biasanya memerlukan data dalam bentuk biner yang menunjukkan keberadaan (1) atau ketidakhadiran (0) suatu item dalam tiap transaksi. Konversi ini mengubah data menjadi lebih mudah diolah oleh algoritma dengan parameter yang dapat disesuaikan, seperti batas nilai untuk konversi atau mapping ulang atribut biner agar sesuai dengan standar input algoritma mining.

Operator FP-Growth merupakan inti dari proses mining pola asosiasi pada workflow ini. FP-Growth termasuk algoritma efisien untuk menemukan frequent itemsets berdasarkan pendekatan pohon FP (Frequent Pattern tree). Parameter utama yang harus disetting pada FP-Growth mencakup minimum support, yang menentukan batas minimum frekuensi kemunculan itemset dalam data untuk dianggap signifikan. Nilai support yang terlalu rendah bisa menyebabkan banyaknya itemset yang kurang berarti, sedangkan nilai yang terlalu tinggi bisa membuat pola penting terlewatkan. Oleh sebab itu, pemilihan nilai support harus disesuaikan dengan karakteristik dataset dan tujuan analisis. Selain itu, parameter maximum pattern length juga biasanya diatur untuk membatasi panjang maksimum deret item dalam pola, agar menghasilkan aturan yang mudah diinterpretasi dan relevan.

Setelah proses mining melalui FP-Growth selesai, hasilnya dialirkan ke operator Create Association Rules. Operator ini mengambil frequent itemsets yang ditemukan dan mengubahnya menjadi aturan asosiasi, yang terdiri dari bagian antecedent (awal) dan consequent (tujuan) dengan nilai support dan confidence sebagai ukuran kekuatan aturan. Parameter penting di tahap ini adalah minimum confidence, yang menetapkan tingkat keyakinan minimum pada aturan yang akan dihasilkan. Confidence menggambarkan seberapa sering aturan yang diajukan benar di dalam data. Pengaturan parameter ini membantu memfilter aturan-aturan yang kurang valid atau kurang signifikan, sehingga hanya aturan dengan kekuatan dan relevansi yang memadai yang ditampilkan.

Keseluruhan alur yang dibangun di RapidMiner ini menunjukkan pendekatan praktis dan terstruktur dalam menjalankan algoritma Apriori, berfokus pada tahap awal pemilihan data, transformasi format data, mining pola dengan FP-Growth, hingga pembuatan aturan asosiasi. Keterpaduan visual drag-and-drop memudahkan pengguna dalam mengatur dan menyesuaikan parameter di tiap operator tanpa memerlukan kode pemrograman yang kompleks, sehingga mempercepat analisis data.

Penting pula untuk dicatat bahwa setiap parameter dalam operator di RapidMiner berperan krusial dalam kualitas hasil. Contohnya, pemilihan atribut yang tepat mencegah noise dan data yang tidak relevan masuk ke dalam analisis, sehingga meningkatkan efektivitas model. Konversi data numerik ke binomial memastikan algoritma Apriori dapat berfungsi optimal karena berdasarkan keberadaan item. Pengaturan support dan confidence secara proporsional menjaga keseimbangan antara pola yang ditemukan banyak dan yang benar-benar berarti, sesuai dengan tujuan riset atau aplikasi analisis.

Selain itu, fitur interaktif RapidMiner memungkinkan validasi hasil secara langsung melalui visualisasi dan laporan yang detail, sehingga pengguna dapat melakukan iterasi pengaturan parameter untuk memperoleh hasil yang paling representatif. Proses iteratif ini sangat mendukung aspek eksplorasi data (exploratory data analysis) yang menjadi salah satu keunggulan platform RapidMiner.

Dengan demikian, langkah teknis menjalankan algoritma Apriori di RapidMiner yang terinci pada workflow ini bukan hanya mengedepankan kemudahan penggunaan, tetapi juga menjunjung tinggi aspek validitas, transparansi, dan efisiensi analisis. Workflow ini sangat sesuai dipakai dalam penelitian yang memerlukan identifikasi pola-pola tersembunyi dalam data transaksi, rekomendasi produk, analisis perilaku konsumen, dan berbagai aplikasi lain yang berorientasi data mining.

Hasil Analisis: Pola-pola Menarik yang Terungkap

Metode Apriori merupakan salah satu teknik populer dalam data mining untuk menemukan pola asosiasi dalam dataset transaksi. Dari hasil yang diperoleh menggunakan RapidMiner, tampak berbagai aturan asosiasi dan frequent itemsets yang terbentuk dengan parameter tertentu yang mencakup nilai support, confidence, LaPlace, dan gain. Analisis ini bertujuan mengidentifikasi pola-pola yang menunjukkan hubungan signifikan antara item antecedent (premises) dengan item consequent (conclusion) dalam dataset transaksi.

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain
156	ATK39	ATK25	0.258	0.571	0.887	-0.845
166	ATK37	ATK25	0.226	0.583	0.884	-0.545
177	ATK42	ATK25	0.323	0.588	0.854	-0.774
180	ATK1	ATK25	0.290	0.600	0.870	-0.671
181	ATK4	ATK25	0.290	0.600	0.870	-0.671
204	ATK17	ATK11, ATK25	0.194	0.600	0.902	-0.455
205	ATK21	ATK11, ATK25	0.194	0.600	0.902	-0.455
206	ATK26	ATK11, ATK25	0.194	0.600	0.902	-0.455
207	ATK7	ATK11, ATK25	0.194	0.600	0.902	-0.455
214	ATK32, ATK6	ATK25	0.194	0.600	0.902	-0.455
217	ATK17	ATK25, ATK32	0.194	0.600	0.902	-0.455
219	ATK21	ATK25, ATK32	0.194	0.600	0.902	-0.455
221	ATK24	ATK25, ATK42	0.194	0.600	0.902	-0.455
222	ATK28, ATK41	ATK25	0.194	0.600	0.902	-0.455
223	ATK24	ATK25, ATK28	0.194	0.600	0.902	-0.455
224	ATK26	ATK25, ATK6	0.194	0.600	0.902	-0.455

Gambar 4 Hasil asosiasi apriori

Pada tabel hasil, setiap baris merepresentasikan sebuah aturan asosiasi yang ditemukan oleh algoritma Apriori. Kolom utama dalam tabel meliputi:

- **Premises:** Ini menunjukkan item atau kombinasi item awal (antecedents) yang muncul dalam transaksi.
- **Conclusion:** Item yang mengikuti dari item awal tersebut (consequents).
- **Support:** Mengukur seberapa sering aturan tersebut muncul dalam dataset, dinyatakan dalam proporsi total transaksi.
- **Confidence:** Mengindikasikan probabilitas bahwa item conclusion muncul jika item premise sudah ada dalam transaksi.
- **LaPlace:** Merupakan metode smoothing probabilitas yang membantu menghindari masalah zero-frequency pada data sparse.
- **Gain:** Merupakan nilai yang mengindikasikan keuntungan relevansi aturan yang ditemukan, dengan nilai negatif menandakan aturan yang mungkin kurang memberikan informasi baru.

Pola Menarik dan Interpretasi Hasil

Dari hasil yang terlihat, beberapa pola menarik dapat diidentifikasi, misalnya aturan seperti “ATK39 → ATK25,” “ATK37 → ATK25,” dan

“ATK42 → ATK25” memiliki nilai support yang cukup tinggi yakni di atas 0,2, dengan confidence sekitar 0,57 hingga 0,60. Hal ini mengindikasikan bahwa produk ATK25 memiliki hubungan yang kuat dan sering muncul bersama dengan produk ATK39, ATK37, dan ATK42 dalam dataset transaksi.

Nilai support yang diperoleh menunjukkan bahwa cukup banyak transaksi yang mengandung kombinasi item-item tersebut. Confidence yang berada di kisaran 0,57-0,60 menunjukkan bahwa jika pelanggan membeli salah satu item antecedent (misalnya ATK39), maka kemungkinan besar mereka juga membeli ATK25 mencapai lebih dari 57%. Hal ini menjadi insight penting bagi strategi pemasaran, misalnya dalam pembuatan paket bundling produk atau pengaturan tata letak toko agar produk-produk ini dapat ditempatkan secara berdekatan untuk meningkatkan penjualan.

Analisis LaPlace dan Gain

Nilai LaPlace yang cenderung mendekati 0,85 hingga 0,90 menunjukkan bahwa aturan-aturan ini memiliki tingkat keandalan yang cukup tinggi dalam mengestimasi probabilitas kemunculan item consequent yang bersangkutan, terutama membantu mengurangi overfitting dengan memperhalus estimasi probabilitas berdasarkan data yang terbatas.

Sebaliknya, skor Gain yang sebagian besar bernilai negatif mengindikasikan bahwa aturan yang ditemukan tidak selalu memberikan “keuntungan informasi” maksimum, atau dengan kata lain, ada kemungkinan beberapa aturan ini kurang efektif dalam memberikan kontribusi informasi baru yang signifikan dibanding model baseline. Oleh karena itu, meskipun beberapa aturan memiliki nilai support dan confidence yang cukup tinggi, perlu dilakukan analisis tambahan untuk memilih aturan yang benar-benar berguna dalam aplikasi nyata.

Kompleksitas Aturan dan Kombinasi Item

Selain aturan dengan single antecedent, terdapat pula aturan yang menggabungkan beberapa item sebagai premise, contohnya seperti “ATK17 \rightarrow ATK11, ATK25” atau “ATK24 \rightarrow ATK25, ATK32.” Adanya aturan dengan premise lebih dari satu item menandakan situasi di mana pembelian kombinasi produk tertentu secara bersamaan meningkatkan peluang munculnya produk lain, yang lebih kompleks namun lebih menggambarkan perilaku konsumen yang realistis.

Keberadaan aturan-aturan kompleks tersebut membuka peluang untuk pemahaman lebih mendalam atas pola pembelian pelanggan terhadap paket produk atau preferensi kombinasi yang sering muncul dalam data transaksi. Ini juga dapat membantu dalam pengembangan strategi cross-selling yang lebih efektif, terutama bagi toko atau perusahaan yang menawarkan banyak jenis produk terkait.

Implikasi Hasil Terhadap Analisis Data dan Pengambilan Keputusan

Hasil analisis dari algoritma Apriori melalui RapidMiner ini memberikan wawasan praktis yang dapat digunakan oleh pelaku bisnis untuk meningkatkan strategi pemasaran dan penjualan produk. Dengan mengetahui kombinasi produk yang sering dibeli bersamaan, pengelola dapat memperkuat kampanye promosi, mengoptimalkan tata letak toko, dan merancang paket produk dengan harga khusus.

Misalnya, jika item ATK25 secara konsisten muncul sebagai consequent dari banyak aturan, maka produk ini dapat dianggap sebagai produk kunci atau produk dengan popularitas tinggi yang mempengaruhi pembelian produk lainnya. Hal ini memperlihatkan pentingnya fokus pada stok dan promosi produk tersebut.

Selain itu, analisis confidence dan LaPlace yang tinggi untuk beberapa aturan memberikan keandalan lebih pada pola yang ditemukan, sehingga memperkuat keputusan bisnis. Namun, perhatian juga harus diberikan

pada nilai gain yang negatif sebagai indikator perlunya evaluasi lebih lanjut atas kebaruan dan nilai tambah dari setiap aturan asosiasi.

Dengan kemudahan dan kecepatan yang ditawarkan software RapidMiner dalam menjalankan algoritma Apriori termasuk dalam mengatur parameter dan mengekstrak hasil, analisis semacam ini menjadi sangat praktis dan efisien baik dalam skala penelitian akademik maupun aplikasi industri.

Hasil Analisis Pola Asosiasi dan Frequent Itemsets Menggunakan Algoritma Apriori

Analisis pola asosiasi merupakan salah satu teknik penting dalam data mining yang bertujuan menemukan hubungan signifikan antar item dalam sebuah dataset transaksi. Dengan menggunakan algoritma Apriori melalui bantuan perangkat lunak RapidMiner, diperoleh berbagai aturan asosiasi dan frequent itemsets yang menggambarkan keterkaitan pembelian produk yang terjadi secara bersamaan. Dataset riset ini kemudian dianalisis lebih lanjut untuk menaikkan wawasan bisnis terutama yang berkaitan dengan perilaku konsumen dan strategi pemasaran.

Struktur Data dan Parameter Kunci

Data hasil analisis terdiri atas beberapa atribut utama untuk setiap aturan asosiasi yang ditemukan, yaitu:

- **Premises:** Item atau kombinasi item awal (antecedents) yang menjadi dasar aturan.
- **Conclusion:** Item atau gabungan item yang muncul sebagai konsekuensi dari premises tersebut.
- **Support:** Frekuensi kemunculan aturan dalam bentuk proporsi dari total transaksi, yang menjadi ukuran popularitas relatif aturan dalam dataset.
- **Confidence:** Probabilitas bahwa conclusion muncul dalam transaksi apabila premises sudah ada. Nilai ini menandakan kekuatan asosiasi.

- LaPlace: Metode smoothing yang membantu memperbaiki estimasi probabilitas agar lebih stabil dan menghindari masalah zero-frequency, penting untuk dataset yang cenderung sparse atau jarang.
- Gain: Nilai yang menunjukkan keuntungan atau penambahan informasi yang diberikan aturan asosiasi dibanding kondisi dasar (baseline).
- Lift: Ukuran yang menentukan kekuatan asosiasi dengan membandingkan frekuensi kemunculan item secara bersama dan secara independen.

Pola Asosiasi Utama dan Implikasinya

Pada data yang dianalisis, terlihat beberapa pola yang cukup dominan dan menarik. Sebagai contoh, aturan seperti:

- $\text{ATK39} \rightarrow \text{ATK25}$ dengan support 0,258 dan confidence 0,57,
- $\text{ATK37} \rightarrow \text{ATK25}$ dengan support 0,226 dan confidence 0,58,
- $\text{ATK42} \rightarrow \text{ATK25}$ dengan support 0,32 dan confidence 0,59,
- serta aturan lain yang secara konsisten menghasilkan conclusion berupa ATK25.

Pola ini memperlihatkan bahwa item produk ATK25 memiliki hubungan yang kuat dan sering ditemukan bersama dengan produk lain seperti ATK39, ATK37, dan ATK42. Dengan tingkat support yang berada di kisaran 20-30%, cukup banyak transaksi mencakup kombinasi tersebut. Confidence di kisaran 0,57 hingga 0,6 menyatakan bahwa peluang pelanggan untuk membeli ATK25 ketika mereka membeli item-item terkait tersebut sangat tinggi, menunjukkan produk ini sebagai produk kunci atau populer.

Kombinasi Premises Lebih Kompleks

Selain aturan-aturan dengan satu item sebagai premises, terdapat juga aturan yang menggabungkan beberapa item sebagai antecedents, seperti:

- $\text{ATK17} \rightarrow \text{ATK11}, \text{ATK25}$

- ATK24 → ATK25, ATK42
- ATK32, ATK6 → ATK25 dan lain-lain

Aturan-aturan ini sekaligus memberikan gambaran perilaku konsumen yang lebih kompleks, di mana pembelian produk dalam bentuk paket atau kombinasi lebih dari satu item dapat memiliki dampak kuat terhadap produk lain. Ini membuka peluang bisnis yang lebih luas dalam merancang strategi paket bundling, penentuan tata letak toko yang strategis, dan pengembangan penawaran cross-selling yang efektif.

Interpretasi Nilai LaPlace, Gain, dan Lift

Nilai LaPlace yang berkisar antara 0,85 hingga 0,90 pada sebagian besar aturan menunjukkan bahwa estimasi probabilitas kemunculan consequent sudah cukup stabil dan dapat diandalkan walaupun data yang digunakan terbatas. Hal ini juga menandai bahwa aturan-aturan ini tidak terlalu dipengaruhi oleh fluktuasi data ekstrim.

Namun, nilai Gain yang dominan negatif pada kebanyakan aturan mengindikasikan bahwa sebagian aturan mungkin tidak memberikan nilai tambah informasi yang signifikan dibanding kondisi dasar. Ini berarti tidak semua aturan dengan nilai support dan confidence tinggi secara otomatis bernilai tinggi secara informatif dari sudut pandang bisnis. Oleh karena itu, analisis lanjutan dan evaluasi konteks aturan diperlukan agar tidak terjadi over-interpretasi data.

Nilai Lift, yang mengindikasikan seberapa jauh associatif relationship lebih kuat dibandingkan dengan hubungan acak, menunjukkan variasi yang cukup besar, dari nilai negatif, mendekati nol, sampai nilai positif kecil. Lift yang positif dan lebih dari 1 menunjukkan adanya asosiasi positif yang bermakna, sedangkan nilai di bawah 1 dapat menunjukkan adanya hubungan negatif atau tidak ada asosiasi.

Wawasan Bisnis dan Strategi Berdasarkan Temuan

Data ini menyediakan sejumlah wawasan strategis berharga, terutama dalam konteks manajemen produk dan penjualan. Produk ATK25 tampak

sangat sentral dalam pola pembelian konsumen, sehingga dapat dijadikan fokus utama dalam penentuan stok, promosi, atau pemasaran produk. Misalnya, toko bisa mengatur produk ATK25 agar mudah diakses bersama produk-produk lain yang berasosiasi secara kuat seperti ATK39, ATK37, dan ATK42.

Strategi bundling produk juga dapat dirancang berdasarkan aturan dengan premises lebih dari satu item, memastikan kombinasi produk yang sukses dalam transaksi dipromosikan sebagai paket khusus dengan harga menarik. Ini tidak hanya meningkatkan penjualan produk tertentu tapi juga mendorong penjualan produk yang berkaitan.

Lebih jauh, analisis ini dapat membantu dalam penentuan pola perilaku pelanggan, segmentasi pasar, dan perencanaan stok yang lebih efisien. Pengambil keputusan dapat menggunakan metrik confidence dan support tertinggi sebagai dasar utama agar fokus hanya pada aturan yang benar-benar relevan dan sering terjadi.

Keterbatasan dan Rekomendasi untuk Analisis Selanjutnya

Meski banyak aturan dan pola ditemukan, penting untuk mencatat adanya keterbatasan seperti:

- Nilai Gain negatif yang menunjukkan perlunya evaluasi ulang aturan agar hanya aturan dengan nilai tambah informasi positif yang diaplikasikan dalam keputusan bisnis.
- Data yang digunakan mungkin memiliki bias pada jenis transaksi tertentu, sehingga generalisasi harus dilakukan hati-hati.
- Tampaknya sebagian aturan memiliki kesamaan nilai confidence, sehingga diversifikasi fitur tambahan seperti sequential pattern atau segmentasi pelanggan dapat memberikan pemahaman lebih komprehensif.

Rekomendasi selanjutnya adalah untuk melakukan validasi aturan dengan data transaksi lebih besar, evaluasi impact aturan pada performa penjualan nyata, serta eksplorasi penggunaan algoritma lanjutan seperti

FP-Growth yang dapat mengefisienkan pencarian itemset frequent dengan kompleksitas lebih besar.

Analisis Apriori yang dijalankan dengan perangkat RapidMiner berhasil mengungkap pola-pola asosiasi menarik yang memperlihatkan hubungan pembelian antar produk secara signifikan. Produk ATK25 menonjol sebagai item sentral yang berasosiasi dengan berbagai produk lain, serta nilai support dan confidence menunjukkan kekuatan dan frekuensi hubungan tersebut.

Meskipun sebagian aturan menunjukkan nilai gain negatif sebagai indikator perlunya evaluasi tambahan, informasi yang diperoleh sangat berguna dalam mendukung pengambilan keputusan bisnis berupa penataan produk, pemasaran terarah, dan penawaran produk paket. Analisis ini menjadi dasar penting untuk meningkatkan efektivitas strategi pemasaran dan pelayanan pelanggan berbasis data transaksi yang konkret dan terukur.

Hasil analisis dari algoritma Apriori melalui RapidMiner ini memberikan wawasan praktis yang dapat digunakan oleh pelaku bisnis untuk meningkatkan strategi pemasaran dan penjualan produk. Dengan mengetahui kombinasi produk yang sering dibeli bersamaan, pengelola dapat memperkuat kampanye promosi, mengoptimalkan tata letak toko, dan merancang paket produk dengan harga khusus.

Misalnya, jika item ATK25 secara konsisten muncul sebagai consequent dari banyak aturan, maka produk ini dapat dianggap sebagai produk kunci atau produk dengan popularitas tinggi yang mempengaruhi pembelian produk lainnya. Hal ini memperlihatkan pentingnya fokus pada stok dan promosi produk tersebut.

Selain itu, analisis confidence dan LaPlace yang tinggi untuk beberapa aturan memberikan keandalan lebih pada pola yang ditemukan, sehingga memperkuat keputusan bisnis. Namun, perhatian juga harus diberikan pada nilai gain yang negatif sebagai indikator perlunya evaluasi lebih lanjut atas kebaruan dan nilai tambah dari setiap aturan asosiasi.

Dengan kemudahan dan kecepatan yang ditawarkan software RapidMiner dalam menjalankan algoritma Apriori termasuk dalam mengatur parameter dan mengekstrak hasil, analisis semacam ini menjadi sangat praktis dan efisien baik dalam skala penelitian akademik maupun aplikasi industry.

Dampak Strategis Berdasarkan Temuan

Analisis data transaksi menggunakan algoritma Apriori telah mengungkap pola-pola asosiasi yang signifikan antara item-item produk yang dibeli pelanggan di Kwek Photo. Temuan ini memiliki implikasi strategis yang penting dalam konteks pemasaran dan manajemen pengelolaan stok barang, yang jika diterapkan dengan tepat dapat meningkatkan efisiensi operasional sekaligus mendorong peningkatan penjualan.

Strategi Pemasaran Berbasis Data

Hasil analisis menunjukkan adanya hubungan kuat antara produk tertentu yang kerap dibeli bersama dalam satu transaksi. Misalnya, item dengan kode ATK25 secara konsisten muncul sebagai consequent dalam banyak aturan asosiasi, dengan confidence berada dalam kisaran 0,57 hingga 0,60 maupun lebih tinggi. Hal ini menandakan bahwa pembelian produk lain seperti ATK39, ATK37, dan ATK42 secara signifikan diasosiasikan dengan pembelian ATK25.

Dari sudut pandang strategi pemasaran, insight ini dapat dimanfaatkan untuk merancang kampanye promosi bundling produk yang menggabungkan ATK25 bersamaan dengan item-item pendukung lainnya. Program diskon paket atau penawaran khusus bagi pembelian kombinasi ini akan mendorong konsumen untuk membeli lebih banyak produk sekaligus, meningkatkan nilai transaksi rata-rata. Selain itu, informasi ini dapat mengarahkan pengelola toko untuk mengatur tata letak produk agar produk-produk yang terkait ditempatkan berdekatan secara fisik, sehingga mempercepat keputusan pembelian impulsif sekaligus meningkatkan visibilitas produk utama.

Lebih jauh, pola pembelian yang melibatkan beberapa item sekaligus (multi-item premises) seperti “ATK17 → ATK11, ATK25” dan “ATK24 → ATK25, ATK42” memberikan peluang untuk melakukan segmentasi pasar yang lebih granular. Promosi yang disesuaikan dengan preferensi cluster pembeli tertentu dapat dikembangkan, sehingga upaya pemasaran menjadi lebih tepat sasaran dan berdampak besar terhadap loyalitas pelanggan.

Optimalisasi Pengelolaan Stok Barang

Temuan dari pola frequent itemsets juga sangat relevan bagi manajemen rantai pasok dan pengelolaan stok. Produk-produk yang sering muncul dalam rules dengan nilai support tinggi, misalnya produk ATK25, sebaiknya menjadi prioritas utama dalam pengaturan stok agar tidak terjadi kekosongan barang yang berpotensi menurunkan kepuasan pelanggan dan peluang penjualan.

Selain itu, rule dengan nilai confidence tinggi menunjukkan ketergantungan pembelian antar produk. Kwek Photo dapat menerapkan strategi stok safety stock atau minimum stock level yang lebih ketat pada produk yang menjadi consequent sering muncul, guna mengantisipasi kebutuhan mendadak akibat permintaan yang meningkat secara bersamaan. Hal ini penting agar stok selalu siap untuk memenuhi permintaan gabungan dari produk yang saling terkait.

Pengelolaan stok yang lebih terpadu ini juga membuka ruang untuk efisiensi dalam pengadaan barang. Dengan mengetahui pola kombinasi produk yang kerap dibeli, Kwek Photo dapat mengoptimalkan pembelian ke pemasok maupun alokasi stok di berbagai titik penjualan sehingga meminimalkan biaya penyimpanan dan risiko kelebihan stok.

Evaluasi Quality of Rules: LaPlace dan Gain

Data metrik tambahan seperti LaPlace dan Gain juga memberikan gambaran kualitas aturan yang ditemukan. Nilai LaPlace yang relatif tinggi mendukung keandalan prediksi probabilitas kemunculan produk consequent, yang memperkuat validitas data untuk diterapkan dalam

keputusan bisnis. Namun, nilai Gain yang umumnya negatif perlu diperhatikan sebagai indikator bahwa beberapa aturan mungkin tidak memberikan informasi baru atau keuntungan strategis tambahan, sehingga seleksi aturan yang benar-benar relevan sangat diperlukan sebelum implementasi strategi.

Peran Teknologi dan Implementasi Sistem Informasi

Pemanfaatan RapidMiner sebagai alat bantu analisis data memberikan nilai tambah berupa kemudahan dan kecepatan proses ekstraksi informasi penting dari dataset besar yang kompleks. Kwek Photo dapat mengintegrasikan hasil analisis ini ke dalam sistem informasi manajemen yang mendukung pengambilan keputusan secara real-time. Contohnya, dashboard monitoring stok yang sinkron dengan pola pembelian dapat membantu manajer operasional dalam pengambilan tindakan cepat dan tepat.

Tantangan dan Solusi Selama Proses

Dalam proses implementasi algoritma Apriori sebagai metode untuk mengekstrak pola-pola asosiasi dari data transaksi Kwek Photo, terdapat sejumlah tantangan utama yang memerlukan perhatian khusus agar analisis dapat berjalan efektif dan hasil yang dihasilkan valid serta bermanfaat. Tantangan-tantangan tersebut mencakup aspek data preprocessing, pemilihan parameter algoritma, serta hambatan teknis dan sumber daya yang muncul selama pemodelan. Mengatasi kendala-kendala ini secara tepat menjadi kunci keberhasilan implementasi dan optimalisasi hasil analisis untuk mendukung strategi bisnis yang lebih cerdas.

Salah satu tantangan utama adalah kualitas dan konsistensi data transaksi yang digunakan sebagai input algoritma Apriori. Data transaksi yang tidak terstruktur dengan baik, mengandung missing values, atau redundansi nilai dapat menyebabkan kesalahan dalam pembentukan frequent itemsets dan aturan asosiasi. Untuk mengatasi hal ini, dilakukan tahap preprocessing data secara menyeluruh yang melibatkan pembersihan

data, standarisasi format atribut, dan transformasi atribut numerik menjadi binomial yang dapat diterima oleh RapidMiner. Proses ini termasuk penggunaan operator seperti “Remap Binominals” yang secara otomatis mengkonversi atribut numerik menjadi biner, sehingga algoritma Apriori dapat mengelola data dengan lebih efisien dan mengurangi noise yang tidak relevan.

Tantangan berikutnya terkait dengan penentuan parameter algoritma Apriori, terutama nilai minimum support dan confidence. Parameter ini sangat menentukan banyaknya aturan asosiasi yang dihasilkan dan kualitasnya. Nilai minimum support yang terlalu rendah mengakibatkan banyaknya aturan yang tidak signifikan sehingga mempersulit analisis dan interpretasi, sedangkan nilai yang terlalu tinggi bisa menyebabkan aturan penting terlewatkan. Begitu pula dengan confidence yang mengindikasikan kekuatan hubungan antar item, harus diatur secara tepat untuk meminimalkan false positives. Dalam proses ini, dilakukan eksperimen iteratif untuk mengoptimalkan nilai parameter, memanfaatkan fitur RapidMiner yang memungkinkan visualisasi dan pengaturan ulang nilai-nilai tersebut secara fleksibel sehingga diperoleh keseimbangan optimum antara kuantitas dan kualitas aturan.

Dari sisi teknis, penggunaan RapidMiner juga menghadirkan sejumlah kendala, seperti pengaturan workflow yang kompleks dan kebutuhan pemahaman mendalam terhadap operator yang tersedia. Kesalahan dalam memilih operator atau urutan proses dapat menyebabkan kegagalan analisis atau hasil yang kurang maksimal. Untuk mengatasi tantangan ini, dilakukan pendalaman dokumentasi penggunaan RapidMiner, serta penggunaan contoh workflow yang telah teruji sebagai referensi. Penyesuaian workflow yang meliputi tahap “Retrieve tabular”, “Select Attributes”, “Remap Binominals”, hingga “FP-Growth” dan “Create Association Rules” dilakukan dengan seksama. Pendekatan ini memastikan pipeline analisis berjalan mulus dan kompatibel dengan spesifikasi data, sekaligus memaksimalkan kecepatan dan efisiensi proses mining.

Selain itu, keterbatasan sumber daya komputasi seperti memori dan prosesor juga menjadi kendala potensial terutama saat bekerja dengan dataset besar yang memerlukan komputasi kompleks. Apriori sendiri dikenal cukup mahal secara komputasi karena eksplorasi frequent itemsets yang ekstensif. Oleh karena itu, solusi yang diterapkan meliputi optimasi dataset dengan seleksi atribut penting dan filtering data minimalisasi noise, serta pemanfaatan metode alternatif seperti FP-Growth yang lebih efisien jika memungkinkan. Dengan demikian, proses mining dapat berjalan dengan lebih ringan dan tidak membebani sistem secara berlebihan, memastikan analisis selesai tepat waktu dan hasil dapat segera digunakan untuk pengambilan keputusan.

Dari segi interpretasi hasil, tantangan muncul dalam memahami dan menerjemahkan nilai metrik seperti support, confidence, lift, dan gain yang dihasilkan algoritma. Beberapa aturan yang muncul mungkin memiliki nilai metrik negatif atau tidak intuitif, sehingga membingungkan dalam pengambilan keputusan strategis. Oleh sebab itu, dilakukan pendampingan interpretasi dengan membandingkan hasil aturan asosiasi terhadap konteks bisnis dan pengetahuan domain Kwek Photo, untuk menentukan mana pola yang relevan dan bernilai guna. Pendekatan analytic storytelling juga dipakai untuk menggambarkan bagaimana pola-pola tersebut bisa diterapkan praktis dalam strategi pemasaran dan pengelolaan stok.

Secara keseluruhan, tantangan yang dihadapi selama pelaksanaan algoritma Apriori dalam riset Kwek Photo ini tidak hanya bersifat teknis, tetapi juga konseptual dan strategi. Solusi yang diterapkan bersifat komprehensif, mencakup aspek preprocessing data, pengaturan parameter, optimasi workflow, pengelolaan sumber daya, hingga pendampingan interpretasi hasil. Langkah-langkah tersebut memastikan bahwa proses mining tidak hanya berjalan lancar, tetapi juga memberikan hasil yang bermakna dan dapat digunakan untuk memajukan bisnis secara signifikan.

BAB 5



Refleksi, Kesimpulan, dan Rekomendasi

Refleksi atas Proses Penelitian

Proses penelitian yang dilakukan dalam studi kasus menggunakan algoritma Apriori untuk menganalisis pola transaksi di Kwek Photo memberikan pengalaman yang sangat berharga dalam memahami dinamika ekstraksi pola asosiasi pada data transaksi nyata. Refleksi atas perjalanan riset ini menunjukkan adanya banyak pelajaran penting, baik dari sisi teknis implementasi, kualitas pengolahan data, hingga interpretasi hasil analisis dalam konteks bisnis yang lebih luas.

Pertama, dari aspek teknis, riset ini mengajarkan betapa krusialnya tahapan preprocessing data dalam memastikan kualitas input algoritma. Dataset mentah sering kali mengandung inkonsistensi, missing values, dan variabilitas format yang jika tidak ditangani dengan baik akan menurunkan akurasi dan relevansi aturan asosiasi yang ditemukan. Penggunaan tools seperti RapidMiner dan operator khusus untuk remapping binominal sangat membantu dalam mengubah data menjadi format yang tepat untuk

algoritma Apriori, sehingga memaksimalkan performa proses mining. Pendalaman pemilihan parameter—seperti nilai minimum support dan confidence—juga menjadi kunci untuk mendapatkan keseimbangan antara aturan yang terlalu banyak (noise) dan aturan yang terlalu sedikit (hilangnya pola penting). Proses trial and error serta pemantauan metrik utama memberi wawasan yang memperkaya pemahaman tentang sifat data dan perilaku algoritma.

Kedua, riset ini juga memberikan refleksi penting terkait kompleksitas pengelolaan workflow analisis data di software RapidMiner. Penggunaan berbagai operator dalam urutan yang tepat diperlukan untuk memastikan kelancaran proses dan akurasi hasil. Kesalahan kecil dalam konfigurasi operator dapat menyebabkan hasil yang bias atau bahkan kegagalan proses. Oleh karena itu, pentingnya dokumentasi proses dan contoh praktik menjadi aspek pembelajaran yang vital untuk penelitian selanjutnya agar proses berjalan lebih efisien dan minim risiko.

Selanjutnya, dari sisi interpretasi hasil, riset ini menunjukkan bahwa data mining bukan sekadar menghasilkan aturan atau pola secara kuantitatif, namun juga memerlukan adanya pendalaman konteks bisnis agar hasil tersebut dapat diterjemahkan menjadi strategi yang bermakna. Misalnya, metrik seperti support, confidence, lift, dan gain perlu diintegrasikan dengan pengetahuan domain untuk menentukan pola mana yang benar-benar bernilai bagi pengambilan keputusan bisnis Kwek Photo. Hal ini menjadi pengingat bahwa kemampuan analisis data harus selalu dilengkapi dengan wawasan bisnis agar rekomendasi yang dihasilkan benar-benar aplikatif.

Selain itu, riset ini juga memunculkan refleksi terkait tantangan keterbatasan sumber daya komputasi, terutama saat memproses dataset yang relatif besar dengan variasi item yang kompleks. Penggunaan algoritma alternatif seperti FP-Growth menjadi solusi strategis untuk mengatasi masalah performa tanpa mengorbankan kualitas hasil mining.

Secara keseluruhan, perjalanan riset ini memperkuat keyakinan bahwa proses data mining adalah siklus iteratif yang menggabungkan

aspek teknis, analytica, dan contextual business understanding secara holistik. Proses pembelajaran yang terjadi selama riset memberikan dasar yang kuat untuk pengembangan riset selanjutnya dengan pendekatan yang lebih matang, serta memperkuat kapabilitas dalam menggunakan alat bantu analisis data modern secara optimal. Refleksi ini sekaligus menjadi pijakan untuk penyempurnaan metodologi dan strategi implementasi yang lebih efektif ke depan.

Kesimpulan Utama yang Mendalam

Dalam penelitian ini, hasil analisis algoritma Apriori pada data transaksi bisnis alat tulis Kwek Photo menyampaikan beberapa temuan signifikan yang memiliki implikasi strategis dan praktis bagi pengembangan usaha. Secara keseluruhan, data mining tersebut mengungkap pola-pola asosiasi yang secara potensial dapat membantu meningkatkan strategi pemasaran, pengelolaan persediaan, dan pelayanan pelanggan.

Pertama, aturan asosiasi yang ditemukan menunjukkan keberadaan hubungan yang relatif kuat antar item spesifik dalam transaksi, meskipun sebagian besar metrik lift dan gain menunjukkan nilai negatif atau rendah. Hal ini menandakan bahwa walaupun pola tersebut sering muncul bersama-sama (support dan confidence cukup tinggi), nilai tambah bisnis yang diperoleh dari asosiasi antar produk perlu ditelaah lebih lanjut, terutama melalui kontekstualisasi dalam model bisnis alat tulis. Dalam kata lain, tidak semua aturan asosiasi secara otomatis layak sebagai insight bisnis bila tidak didukung oleh analisis nilai ekonomis dan pola pembelian pelanggan secara menyeluruh.

Kedua, temuan ini menggarisbawahi pentingnya pendekatan hibrid yang mengintegrasikan teknik penggalian data dengan pengetahuan domain bisnis yang mendalam. Aspek interpretasi hasil mining harus terhubung langsung pada tujuan strategis seperti identifikasi produk komplementer yang bisa didorong dalam promosi bundling, pengurangan stok berlebih dengan melihat pola pembelian bersama yang kurang efisien,

serta pengoptimalan variasi produk yang sesuai kebutuhan pasar. Tanpa penyesuaian ini, hasil mining cenderung bersifat deskriptif dan kurang aplikatif.

Ketiga, prosedur analisis yang dilakukan juga mencerminkan kebutuhan akan proses penyesuaian parameter algoritma yang cermat dan trial & error untuk mencapai keseimbangan antara aturan yang representatif dan noise. Penggunaan software RapidMiner sebagai tool yang memudahkan workflow mining secara visual dan modular menjadi kekuatan dalam mempercepat proses, namun tetap memerlukan ketelitian dalam pemilihan nilai ambang batas support dan confidence sesuai karakteristik data transaksi alat tulis yang beragam.

Keempat, hasil riset ini memberikan pijakan kuat bagi Kwek Photo untuk mengadopsi strategi berbasis data. Meskipun tantangan dalam hal interpretasi metrik seperti lift negatif dan gain yang rendah memerlukan perhatian khusus, pemahaman pola asosiasi ini membuka peluang bagi perencanaan pemasaran yang lebih fokus dan efisien, serta peningkatan pengalaman belanja pelanggan melalui rekomendasi produk yang relevan. Dengan demikian, data mining tidak hanya menjadi alat analitik, melainkan fondasi dalam pengembangan bisnis berbasis insight-data.

Secara keseluruhan, kesimpulan ini menunjukkan bahwa pemanfaatan algoritma Apriori pada bisnis alat tulis memberikan nilai strategis dalam menemukan pola tersembunyi yang dapat ditindaklanjuti, asalkan hasil analisis selalu diselaraskan dengan tujuan bisnis dan evaluasi menyeluruh dari segi manfaat ekonomis dan operasional.

Saran untuk Pengembangan Berikutnya

Dalam penelitian ini, hasil analisis algoritma Apriori pada data transaksi Kwek Photo memberikan gambaran penting untuk pengembangan riset dan aplikasi bisnis di masa depan. Pertama, disarankan agar penelitian lanjutan mengadopsi pendekatan multi-algoritma atau hybrid yang tidak hanya mengandalkan Apriori, tetapi juga mengeksplorasi teknik

lain seperti FP-Growth, Eclat, atau metode klasifikasi berbasis machine learning untuk mengungkap pola yang lebih kompleks dan beragam dari segi asosiasi maupun prediksi perilaku konsumen. Hal ini akan membantu memperoleh insight yang lebih robust dan adaptif terhadap dinamika pasar yang terus berubah.

Kedua, pengembangan aplikasi di masa depan sebaiknya fokus pada integrasi analisis data transaksi dengan sistem manajemen persediaan real-time dan pemasaran digital yang memungkinkan pemanfaatan aturan asosiasi untuk rekomendasi produk otomatis, personalisasi promosi, dan optimasi stok secara lebih responsif. Penerapan dashboard interaktif berbasis data mining menjadi sebuah solusi inovatif agar pengambil keputusan dapat dengan cepat memahami pola dan mengambil tindakan strategis yang tepat.

Ketiga, penting pula untuk memperluas cakupan data yang dianalisis, termasuk data demografis pelanggan dan variabel eksternal seperti musim, tren pasar, atau kegiatan promosi agar hasil analisis menjadi lebih kontekstual dan akurat dalam menyesuaikan strategi bisnis.

Akhirnya, penelitian berikutnya dianjurkan untuk mengembangkan framework evaluasi menyeluruh yang tidak hanya mengandalkan metrik klasik (support, confidence, lift, gain), tetapi juga mempertimbangkan dampak bisnis nyata seperti profitabilitas, kepuasan pelanggan, dan daya saing produk.

Tinjauan Manfaat Praktis dan Kebijakan

Hasil analisis data transaksi menggunakan algoritma Apriori memberikan manfaat praktis yang signifikan bagi pemilik usaha alat tulis Kwek Photo serta bagi pengambil kebijakan bisnis. Dari segi praktis, hasil riset ini memungkinkan pemilik usaha untuk mengidentifikasi pola pembelian konsumen secara lebih mendalam, khususnya kombinasi produk yang sering dibeli bersama. Dengan memahami hubungan ini, manajemen dapat mengoptimalkan pengaturan stok sehingga lebih efisien

dan responsif terhadap kebutuhan pasar, mengurangi risiko kelebihan atau kekurangan persediaan yang merugikan.

Selain itu, temuan asosiasi produk juga menjadi dasar strategis untuk merancang program promosi yang lebih efektif dan tepat sasaran, seperti bundling produk yang populer berdasar data transaksi riil. Ini berpotensi meningkatkan nilai penjualan per transaksi sekaligus meningkatkan loyalitas pelanggan melalui rekomendasi produk yang relevan.

Dari sisi kebijakan, pemilik usaha dapat menggunakan hasil riset ini sebagai alat bantu pengambilan keputusan berbasis data (data-driven decision-making). Kebijakan pengelolaan produk dan promosi dapat didesain agar lebih adaptif mengikuti perilaku pembelian aktual dan tren pasar nyata, bukan hanya berdasarkan intuisi atau asumsi. Misalnya, penyesuaian harga dan promo musiman bisa dilakukan secara dinamis sesuai pola asosiasi yang ditemukan.

Secara keseluruhan, implementasi hasil riset ini mendukung transformasi digital bisnis Kwek Photo menjadi lebih berorientasi data dan tepat sasaran, memberikan fondasi kuat bagi pengembangan usaha berkelanjutan dalam menghadapi persaingan pasar yang kompetitif.

Daftar Pustaka

- Y. Liu, "Application and Practice of Data Mining Techniques," *Appl. Comput. Eng.*, vol. 4, no. 1, pp. 237–241, 2023, doi: 10.54254/2755-2721/4/20230456.
- W. Wu *et al.*, "Data Mining in Clinical Big Data: The Frequently Used Databases, Steps, and Methodological Models," *Mil. Med. Res.*, vol. 8, no. 1, 2021, doi: 10.1186/s40779-021-00338-z.
- S. Ninoria and S. S. Thakur, "An Efficient Algorithm for Mining High Utility Rare Itemsets Over Uncertain Databases," *Int. J. Comput. Eng. Technol.*, vol. 10, no. 2, 2019, doi: 10.34218/ijcet.10.2.2019.012.
- M. M. Shamie and M. M. Almustafa, "Improving Association Rule Mining Using Clustering-Based Data Mining Model for Traffic Accidents," *Rev. Comput. Eng. Stud.*, vol. 8, no. 3, pp. 65–70, 2021, doi: 10.18280/rces.080301.
- J. Heikal, "Enhancing Retail Supermarket Financial Performance Through Market Basket Analytics Using Apriori Algorithm in Indonesia Market Case," *Appl. Quant. Anal.*, vol. 4, no. 1, pp. 42–53, 2024, doi: 10.31098/quant.2153.
- A. K. AL-Mashanji, A. H. Hamza, and L. H. Alhasnawy, "Computational Prediction Algorithms and Tools Used in Educational Data Mining: A Review," *J. Univ. Babylon Pure Appl. Sci.*, pp. 87–99, 2023, doi: 10.29196/jubpas.v31i1.4531.
- Q. Bu, "Editorial: Application of Data Mining in Pharmaceutical Research," *Front. Pharmacol.*, vol. 15, 2024, doi: 10.3389/fphar.2024.1388738.
- [8] R. Takdirillah, "Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan," *Edumatic J. Pendidik. Inform.*, vol. 4, no. 1, pp. 37–46, 2020, doi: 10.29408/edumatic.v4i1.2081.

- [9] A. Giovanni, D. J. Pane, and E. Afandi, "Implementasi Metode Association Rule Menggunakan Algoritma Apriori Pada Data Transaksi Penjualan," *J. Sist. Inf. Triguna Dharma (Jursi Tgd)*, vol. 1, no. 3, p. 149, 2022, doi: 10.53513/jursi.v1i3.5116.
- [10] R. K. M, J. Venkatesh, and A. Rahman, "Data Mining and Machine Learning in Retail Business: Developing Efficiencies for Better Customer Retention," *J. Ambient Intell. Humaniz. Comput.*, 2021, doi: 10.1007/s12652-020-02711-7.
- [11] R. Saputra and A. J. Sibarani, "Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat," *Jatisi (Jurnal Tek. Inform. Dan Sist. Informasi)*, vol. 7, no. 2, pp. 262–276, 2020, doi: 10.35957/jatisi.v7i2.195.

Riwayat Penulis



Anggi Ervianti

Lahir di Rantauprapat pada tanggal 15 April 2002. Sejak kecil, saya menempuh pendidikan dasar hingga menengah di kota kelahiran saya. Setelah itu, saya sempat melanjutkan studi sebagai mahasiswa program Sarjana (S1) di Universitas Labuhanbatu dengan jurusan Sistem Informasi. Di jurusan ini, saya mempelajari

keterkaitan antara teknologi informasi dan kebutuhan bisnis, termasuk pengembangan sistem informasi, pengelolaan data, serta pemanfaatan teknologi digital untuk pengambilan keputusan yang lebih efektif. Saya memiliki kemampuan dalam mengoperasikan Microsoft Office dan tertarik dalam dunia desain grafis. Ketertarikan tersebut saya salurkan dengan mengembangkan keterampilan dalam menggunakan berbagai aplikasi desain seperti Adobe Photoshop, Canva, dan CorelDRAW. Bagi saya, desain bukan hanya soal estetika, melainkan juga cara menyampaikan informasi secara visual dengan cara yang lebih menarik dan mudah dipahami.

Tujuan saya menerbitkan buku dengan judul “Penerapan Data Mining dalam Menentukan Pola Penjualan Alat Tulis Kantor di Kwek Photo Menggunakan Metode Apriori” adalah untuk membagikan pengetahuan dan pengalaman saya dalam mengolah data penjualan agar dapat memberikan insight yang bermanfaat bagi pelaku usaha, khususnya di bidang ritel alat tulis kantor. Melalui penerapan metode Apriori dalam data mining, buku ini diharapkan dapat membantu pembaca memahami bagaimana

pola pembelian konsumen dapat dianalisis untuk meningkatkan strategi penjualan, efisiensi stok barang, serta pengambilan keputusan berbasis data. Manfaat dari penelitian ini tidak hanya bersifat akademis, tetapi juga aplikatif bagi pelaku usaha yang ingin mengembangkan bisnisnya secara lebih terarah dan berbasis teknologi informasi.



Ibnu Rasyid Munthe, S.T., M.Kom.

Lahir di Kota Rantau Prapat, Kabupaten Labuhanbatu, pada tahun 1987. Beliau memperoleh gelar Sarjana Teknik (S.T.) dari Universitas Nurtanio Bandung dan melanjutkan pendidikan Magister di Universitas Putra Indonesia (YPTK) Padang dengan gelar Magister Komputer (M.Kom.). Beliau aktif dalam menulis artikel ilmiah serta buku akademik. Saat ini, beliau berkiprah sebagai dosen di Universitas Labuhanbatu. Selain itu, beliau juga mengelola kanal YouTube @manjaddawajada2022 sebagai media pembelajaran dan sarana berbagi ilmu pengetahuan.



Angga Putra Juledi, S.Kom., M. Kom.

Lahir di Kota Padang pada tanggal 19 Juli 1994. Dalam menempuh Pendidikan dimulai dari Sekolah Dasar SDN 19 Padang tamat tahun 2006, SMPN 3 Padang tamat pada tahun 2009, dan di SMA Pertiwi 2 Padang tamat pada tahun 2012. Lalu melanjutkan ke pendidikan perguruan tinggi swasta yaitu S1 (Sarjana) Universitas Putra Indonesia "YPTK" Padang lulus pada tahun 2018 dengan jurusan Sistem Informasi, Dan melanjutkan Program Pascasarjana (S2) di Universitas Putra Indonesia "YPTK" Padang pada tahun 2019 Program Studi Teknik Informatika. Konsentrasi Sistem Informasi.

Saya mengabdikan diri sebagai salah satu Dosen di bidang Ilmu Komputer pada Fakultas Sains Dan Teknologi dengan Program Studi Sistem Informasi di Universitas LabuhanBatu dan menjadi dosen tetap pada tahun 2020 pada kampus tersebut. Saat ini menjadi bagian Struktural di Universitas Labuhanbatu sebagai Kepala Bagian Sumber Daya Manusia periode 2023 s/d 2027. Buku pertama terbit pada 31 Desember 2021 dengan judul Internetworking Dan TCP/IP. Buku kedua terbit pada tanggal 17 Oktober 2023 dengan judul Panduan Belajar HTML, CSS, dan JavaScript untuk Pemula. Hingga saat ini masih menulis buku setiap tahunnya.



Marnis Nasution, S.Kom., M.Kom.

Lahir di Bengkulu 30 maret 1990. Selama sekolah dasar sampai menengah ditempuh di kota Bengkulu. Melanjutkan Pendidikan tinggi strata-1 dan strata-2 di Universitas Putra Indonesia "YPTK" Padang dari tahun 2008 sampai 2024 dengan jurusan Sistem Informasi. Saat ini aktif menjadi Dosen Yayasan di Universitas Labuhanbatu, Sumatera Utara dan menulis beberapa karya Ilmiah dan buku.



Penerapan Data Mining

dalam

Menentukan Pola Penjualan

Alat Tulis Kantor di Kwek Photo Menggunakan Algoritma Apriori

Dalam era digital dan transformasi teknologi informasi yang pesat, pengolahan data besar (big data) telah menjadi aspek vital dalam pengambilan keputusan bisnis yang efektif dan efisien. Salah satu pendekatan yang semakin banyak diaplikasikan dalam dunia bisnis adalah data mining, yaitu teknik dan metode untuk menggali informasi tersembunyi dan pola-pola berharga dari kumpulan data yang sangat besar. Buku monograf ini hadir dengan tujuan memberikan pemahaman mendalam mengenai implementasi data mining, khususnya algoritma Apriori, dalam konteks menentukan pola penjualan alat tulis kantor di usaha ritel Kwek Photo.

Kajian ini difokuskan pada eksplorasi pola asosiasi antar produk yang dapat membantu manajemen dalam memahami perilaku konsumen serta dinamika pasar. Algoritma Apriori, sebagai salah satu metode asosiasi rule mining yang populer, dipilih karena kemampuannya dalam menemukan hubungan signifikan antar item yang sering muncul bersama dalam transaksi penjualan. Dengan penerapan algoritma ini, diharapkan diperoleh insight strategis untuk pengelolaan stok, penentuan promosi, dan peningkatan layanan yang berorientasi pada kebutuhan nyata pelanggan.

Selain aspek teknis, buku ini juga membahas secara komprehensif dampak strategis, tantangan yang dihadapi selama proses analisis data, serta rekomendasi praktis dan kebijakan yang relevan bagi pengembangan usaha Kwek Photo ke depan. Melalui pendekatan empiris dan berbasis data, diharapkan pembaca, terutama para praktisi dan akademisi di bidang manajemen bisnis dan teknologi informasi, dapat memperoleh pemahaman yang utuh dan aplikatif mengenai penerapan data mining dalam praktik nyata.



✉ literasinusantaraofficial@gmail.com
🌐 www.penerbitlitrus.co.id
📘 Literasi Nusantara
📱 [literasinusantara_](https://www.instagram.com/literasinusantara_)
☎ 085755971589

Teknologi

+17

ISBN 978-634-234-400-2



9 786342 344002