

ANALISIS POLA PEMBELIAN KONSUMEN TERHADAP PRODUKS OPERATOR TELEKOMUNIKASI MENGGUNAKAN ALGORITMA C4.5 (STUDI KASUS : FEEYZA CELL)

Dwi Prasetyo, Nana Suarna, Tati Suprapti

Teknik Informatika, STMIK IKMI Cirebon

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

dprast39@gmail.com

ABSTRAK

Kemajuan teknologi komunikasi dan informasi yang sangat pesat dan semakin hemat biaya telah memungkinkan masyarakat untuk beralih menggunakan kuota dan pulsa sehingga kini menjadi suatu kebutuhan yang mutlak. Penulis mulai tertarik. Berbagai produk operator yang tersedia di counter Feezya Cell, seperti Kartu Telkomsel, Kartu XL, Kartu Indosat, Kartu 3, dll, telah dikategorikan oleh peneliti sebagai 'best seller' dan 'non-best seller'. Tujuannya adalah untuk mengetahui penerapan data mining dengan menggunakan algoritma C4.5 dalam menentukan keakuratan data klasifikasi produk terlaris dan hasil persediaan penjualan kartu perdana. Dengan demikian nilai entropy dan gain tertinggi yaitu 2,310770059 pada atribut-atribut Jenis Kartu Perdana pada perhitungan manual. Sedangkan menggunakan aplikasi Rapidminer diperoleh hasil pohon keputusan seperti terlihat pada Gambar 4.8 Harga – Jumlah Terjual – Jenis Kartu Perdana (3 = Kurang Laris, Telkomsel = Laris) jika harganya di atas 5500 dan kurang dari atau sama dengan 6500, disarankan Indosat; jika di atas 6500 dan kurang dari atau sama dengan 7500, disarankan Telkomsel atau XL berdasarkan kondisi tertentu.

Kata kunci : Produk Telekomunikasi, Algoritma C4.5, Data Mining, Klasifikasi

1. PENDAHULUAN

Komunikasi merupakan kebutuhan universal semua orang. Tanpa komunikasi, orang tidak bisa mengenal satu sama lain dengan lebih baik. Meskipun setiap orang di bumi berbicara dalam bahasa yang berbeda, komunikasi adalah cara yang efektif untuk berinteraksi dengan orang lain. Teknologi komunikasi saat ini berkembang pesat, memungkinkan masyarakat untuk mengirim dan menerima pesan dengan cepat, mudah, dan fleksibel. Dahulu, pesan dikirim dan diterima melalui surat, telegram, faks, dan telepon rumah, namun kini telah menjadi media komunikasi yang sangat nyaman yang memungkinkan pesan dikirim dan diterima melalui suara, teks, dan gambar kapan saja dan di mana saja. Teknologi komunikasi ini memungkinkan komunikasi hanya dalam cakupan area tertentu. Teknologi yang dimaksud adalah telepon seluler.

Permasalahan dalam penelitian ini adalah ketatnya persaingan antar operator telepon seluler untuk mempertahankan dan memperoleh pelanggan baru yang menguntungkan masyarakat. Konsumen mendapat manfaat dari perusahaan yang mengenakan biaya, karena operator jaringan menurunkan biaya untuk mempertahankan pelanggan dan menarik pelanggan baru. Harga Kartu Perdana yang murah memungkinkan konsumen untuk memilih produk telepon seluler, sehingga konsumen sering berganti-ganti Kartu Perdana, dan sebagian konsumen menggunakan beberapa merek penyedia.

Penelitian ini bertujuan untuk menganalisis perilaku pembelian konsumen terhadap produk operator telekomunikasi dengan menggunakan metode algoritma C4.5. Penelitian ini bertujuan untuk

memberikan wawasan tentang perilaku pembelian konsumen dan faktor-faktor yang mempengaruhi keputusan pembelian konsumen berdasarkan data historis.

2. TINJAUAN PUSTAKA

2.1. Pengertian Data Mining

Data mining adalah penemuan pola, pengetahuan, dan wawasan berguna yang tersembunyi dalam data yang besar dan kompleks. Ini melibatkan penggalian informasi berguna dari kumpulan data besar dengan tujuan mengidentifikasi pola yang dapat digunakan untuk membuat keputusan, membuat prediksi, atau menemukan pengetahuan baru. Penambangan data menggunakan berbagai teknik dan algoritma seperti clustering, klasifikasi, regresi, asosiasi, dan visualisasi data untuk mengungkap pola yang tidak langsung terlihat oleh manusia. Tujuan utamanya adalah mengubah data mentah menjadi informasi yang berguna dan dapat diterapkan.

2.2. Pengertian Klasifikasi Prediksi

Klasifikasi prediksi melibatkan prediksi kelas atau label suatu objek berdasarkan atributnya. Ini adalah salah satu tugas utama data mining dan pembelajaran mesin. Dalam konteks ini, "klasifikasi" mengacu pada pengelompokan data ke dalam kelas atau kategori yang berbeda, dan "prediksi" mengacu pada kemampuan model untuk memprediksi kelas data baru dan yang tidak diketahui.

2.3. Pengertian AI Goritma C4.5

Membangun pohon keputusan, yang merupakan model prediktif dalam data mining dan pembelajaran mesin. Algoritma ini dikembangkan oleh Ross Quinlan sebagai pengembangan dari algoritma sebelumnya, ID3 (Iterative Dichotomiser 3). Setelah penelitian melakukan telaah beberapa penelitian, ada beberapa yang memiliki keterkaitan dengan penelitian yang peneliti lakukan. Penelitian yang berhasil peneliti temukan adalah Berdasarkan hasil Literature review dan perbandingan algoritma atau skema yang diterapkan pada beberapa paper di bawah ini :

Pada penelitian mengkategorikan berbagai produk pulsa yang tersedia di counter RA Cell – Pulsa Telkomsel, Pulsa XL, Pulsa Indosat dan 3 Pulsa menjadi ‘hotselling’ dan ‘unselling’. Tujuannya adalah untuk mengetahui penerapan data mining menggunakan algoritma Naive Bayes dalam menentukan hasil keakuratan data untuk klasifikasi produk terlaris dan persediaan penjualan pulsa. Mengumpulkan 600 data menjadi 480 data dan 120 data uji. Data mining adalah suatu bentuk penambangan data yang menggunakan tool RapidMiner dan algoritma Naive Bayes untuk mengklasifikasikan data dalam jumlah besar.[1]

Adapun Penelitian bertujuan untuk mengetahui Pengaruh Citra Merek Terhadap Keputusan Pembelian Kartu Perdana Tri Pada Warga Kelurahan Pondok Karya Kota Tangerang Selatan. Metode penelitian yang digunakan yaitu deskriptif kualitatif dengan pemilihan sampel random. Hasil penelitian menunjukkan bahwa citra merek berpengaruh positif terhadap keputusan pembelian dengan persamaan regresi $Y = 25,037 + 2,222X$. Semakin baik citra merek maka keputusan pembelian akan semakin meningkat dan tingkat hubungan 0,379 yang “rendah” dengan kontribusi pengaruh sebesar 14,4%. Hasil uji t diperoleh thitung 11,267 > ttabel 1,652 dan signifikansi 0,000. [2]

Penelitian ini bertujuan untuk menemukan pengaruh antara harga terhadap keputusan pembelian. Populasi penelitian berjumlah 80 orang yang seluruhnya diteliti. Teknik sampling yang digunakan adalah teknik Sampling Insidental. Setelah data dinyatakan berdistribusi normal dan linier berdasarkan uji normalitas dan uji linearitas, dilanjutkan dengan analisis varians (ANOVA). Berdasarkan hasil perhitungan regresi linier diperoleh $Y = 51,10 + 0,477 x$ dan hasil perhitungan ANOVA didapat Fhitung < Ftabel yaitu 0,732 < 1,75. Pengujian hipotesis dilakukan dengan menghitung korelasi product moment dan didapat Rhitung > Rtable 0,381 > 0,220 dan Thitung > Ttabel 4,67 > 1,66 yang berarti H_0 ditolak dan H_a diterima. Oleh karena itu hasil penelitian ini disimpulkan bahwa ada pengaruh antara Harga terhadap Keputusan Pembelian pada Konsumen Kartu Seluler Telkomsel.[3]

Tujuan dari penelitian ini adalah dilakukan tanpa menerapkan metode untuk memprediksi akan kebutuhan pakaian syar’i diminati oleh konsumen berdasarkan data penjualan bulan-bulan sebelumnya. Oleh karena belum diterapkannya metode prediksi pola pembelian konsumen pada Toko Rabbani Urip, mengakibatkan pihak toko sering membeli/memesan produk syar’i yang kurang diminati konsumen ke pabrik atau supplier. Dengan adanya pemesanan produk pakaian syar’i yang kurang diminati oleh konsumen dapat menyebabkan tumbukan pakaian di gudang yang tentunya akan menambah biaya persediaan serta dapat merugikan pihak toko pakaian syar’i. Penelitian ini bertujuan untuk membuat dan mengimplementasikan algoritma C.45 dalam sebuah aplikasi penentuan pola penjualan pada Toko Rabbani Makassar. Sistem yang dirancang terdiri dari 22 rancangan input, 7 rancangan tabel basis data, dan menghasilkan 7 rancangan output. Berdasarkan hasil pengujian black-box testing terhadap pengujian black-box terhadap ke-21 modul dalam sistem pengenalan nota penjualan Toko Rabbani Urip Makassar menggunakan algoritma C.45 maka disimpulkan bahwa sistem secara keseluruhan telah bekerja sesuai dengan yang diharapkan.[4]

Penelitian ini memiliki banyaknya pengguna Telkom genggam/ gadget di dunia teknologi telekomunikasi khususnya bidang seluler memberi banyak kemudahan bagi pelanggan untuk melakukan transaksi pembelian pulsa elektronik. Persaingan antar supplier pulsa elektrik membutuhkan suatu analisa untuk mengetahui tingkat kepuasan pelanggan pulsa elektronik. Penelitian ini menggunakan Algoritma C4.5 dan C4.5 Berbasis Bagging yang akan digunakan dalam menganalisa kepuasan pelanggan, dimana data taining menggunakan 460 pelanggan pulsa elektronik serta menggunakan 11 variabel yang digunakan. Penelitian ini bertujuan menghasilkan analisa kepuasan konsumen menggunakan algoritma C4.5 dan C4.5 berbasis bagging untuk mengetahui algoritma yang sesuai dalam menganalisa kepuasan konsumen pulsa elektronik. [5]

Penelitian ini tujuan perusahaan dalam memberikan kepuasan kepada konsumen baik perusahaan jasa maupun non jasa. Salah satu perusahaan penyedia layanan akses internet yang berkomitmen meningkatkan kualitas layanan terhadap kepuasan pelanggan adalah Hinet Batam. Penelitian ini bertujuan untuk menganalisis kualitas pelayanan terhadap kepuasan konsumen Hinet Batam dengan menggunakan algoritma C4.5. Variabel penilaian meliputi empati, daya tanggap, keandalan dengan keputusan puas dan tidak puas. Hasil pengolahan metode Data Mining dengan algoritma C4.5 mampu mengklasifikasikan tingkat kepuasan pelanggan terhadap kualitas layanan. Dengan nilai presisi prediksi yang memuaskan sebesar 61,7% dan nilai presisi prediksi yang tidak memuaskan sebesar 38,3%. Hasil penelitian ini dapat dijadikan acuan

untuk meningkatkan kualitas pelayanan dan menjaga loyalitas pelanggan dalam menggunakan produk paket data layanan internet serta meningkatkan penjualan. [6]

Penelitian ini adalah Pemenuhan kebutuhan dan keinginan pelanggan merupakan salah satu strategi dalam mempertahankan pelanggan. Diantara strategi tersebut adalah kualitas produk dan pelayanan yang berkualitas juga akan memenuhi harapan pelanggan. Tujuan dari penelitian ini adalah untuk memprediksi apakah pelanggan akan menjadi pelanggan setia atau meninggalkan layanan perusahaan. Metode penelitian menggunakan teknik CRISP-DM (Cross Industry Standard Process for Data Mining) dan metode tuning hypermeter dengan algoritma ridge classifier dan matriks konfusi. CRISP-DM merupakan model proses yang menjadi dasar proses ilmu data dengan enam fase berurutan. Data yang digunakan sebanyak 7.043 record. dibagi menjadi 70:30 data train dan data test. Dari pilihan fitur, sebagian besar pelanggan yang berhenti berlangganan adalah pelanggan yang tidak menggunakan beberapa layanan seperti VPN, Cadangan Data, Perlindungan Perangkat, Bantuan Teknisi, Streaming TV, Streaming Film, kontrak bulanan dan menggunakan metode pembayaran E-Wallet. Kemudian hasil penelitian akurasi dengan menggunakan Confusion Matrix menunjukkan hasil yang cukup baik dengan akurasi sebesar 80,5%, Precision sebesar 85,7% dan Recall sebesar 89,8%. Sebanyak 73% atau 5.174 tetap menggunakan layanan tersebut dan 1.869 atau sekitar 27% pelanggan berhenti menggunakan layanan Telkom Indonesia. [7]

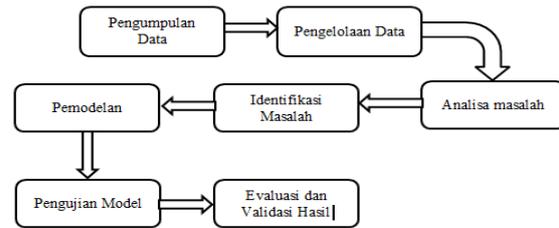
Penelitian ini bertujuan untuk kepuasan konsumen merupakan tingkat dimana anggapan terhadap produk sesuai dengan harapan para konsumen. Harapan konsumen umumnya merupakan prakiraan atau keyakinan konsumen tentang apa yang akan diterimanya bila telah membeli atau mengkonsumsikan suatu produk. Kenyataannya, apa yang bisa memuaskan konsumen di satu situasi mungkin tidak bisa memuaskan konsumen yang sama di lain situasi. Ditambah dengan sistem keluhan konsumen yang hanya disediakan contact person oleh pihak restoran. Dengan demikian, kepuasan konsumen sangat sulit diketahui, karena tidak semua konsumen menghubungi contact person yang disediakan. Maka dari itu, adanya penelitian mengenai analisis kepuasan konsumen di Restoran Dapur Solo (Cabang Kediri) tersebut diperlukan. Untuk memenuhi kebutuhan-kebutuhan diatas, banyak carayang dapat ditempuh. Salah satunya adalah dengan melakukan pemanfaatan data perusahaan yang diolah dengan klasifikasi Decision Tree menggunakan algoritma C4.5. Dari proses klasifikasi akan menghasilkan beberapa aturan yang menyebabkan konsumen merasa puas dan tidak puas, serta seberapa besar konsumen merasa puas maupun tidak puas. [8]

Penelitian ini bertujuan untuk mengetahui pengaruh citra merek dan harga terhadap keputusan pembelian kembali kartu seluler Telkomsel terhadap konsumen di Kutai Kartanegara Muara Jawa. Metode penelitian yang digunakan adalah kuantitatif. Subjek dalam penelitian ini adalah 100 orang. Metode pengumpulan data menggunakan tiga skala yaitu skala keputusan pembelian kembali, skala citra merek dan skala harga dengan model skala Likert. Sampel penelitian diambil dengan menggunakan teknik purposive sampling. Data yang terkumpul dianalisis dengan model bertahap dan uji regresi model penuh dengan bantuan program Statistik Paket untuk Ilmu Sosial (SPSS) 20.0 untuk Windows. Dari analisis data, hasil yang diperoleh untuk model penelitian ini menunjukkan bahwa ada pengaruh antara citra merek dan keputusan pembelian berulang, nilai yang diperoleh adalah $t_{hitung} = 2,080 > t_{tabel} = 1,984$ dan $p = 0,040 < 0,050$. Hasil analisis kedua menunjukkan pengaruh harga terhadap keputusan pembelian berulang, nilai yang diperoleh adalah $t_{hitung} = 6,313 > t_{tabel} = 1,984$ dan $p = 0,000 < 0,050$. Hasil analisis ketiga menunjukkan bahwa terdapat pengaruh antara citra merek dan harga terhadap keputusan pembelian kembali kartu seluler Telkomsel kepada konsumen di Kutai Kartanegara Jawa, nilai yang diperoleh adalah $F_{hitung} = 40,112 > F_{tabel} = 3,09$, $R^2 = 0,453$, dan $p = 0,000 < 0,050$. Hasil yang ditunjukkan dalam penelitian ini adalah bahwa keputusan pembelian kembali dipengaruhi secara positif oleh citra merek dan harga pada konsumen yang membeli kembali produk dari kartu seluler Telkomsel di Kutai Kartanegara Muara Jawa. [9]

Penelitian ponsel ini merupakan sebuah bisnis seluler yang bergerak dalam bidang penjualan kartu perdana internet. Kartu Perdana Internet adalah kartu yang digunakan dengan pembatasan atau batasan penggunaan internet, dimana dengan menggunakan kartu perdana internet ini pengguna dapat menghemat pulsa karena semua aktivitas menggunakan kuota kecuali telepon (panggilan suara) dan Short Message Service (SMS). Beberapa perusahaan bersaing dengan mengeluarkan internet premiere sehingga dapat menekan biaya akses. Persaingan ini terlihat dari banyaknya kartu perdana internet yang terjual secara bersamaan dan mengetahui persediaan untuk bulan berikutnya. Maka untuk memudahkan perhitungan pemilik ponsel menggunakan data mining dengan metode algoritma apriori untuk menyelesaikan penelitian ini. Dengan menggunakan metode algoritma apriori dapat membantu pemilik ponsel purnama dalam proses perhitungan pada ponsel purnama menggunakan aplikasi Tanagra dan diharapkan dapat mengurangi jumlah kartu yang kadaluarsa serta dapat memprediksi berapa persediaan kartu yang dibutuhkan pada bulan berikutnya. [10]

3. METODE PENELITIAN

Tahapan yang akan digunakan dalam penelitian ini dapat dilihat pada table 3. 1 untuk melakukan Klaasifikasi Prediksi terhadap data hasil Pembelian Produk operator telekomunikasi, berikut ini beberapa langkah dalam tahapan penelitian yang akan dilakukan.



Gambar 1. Metode Penelitian

Tabel 1. Deskripsi Metode Penelitian

Tahapan	Deskrip Aktivitas
Pengumpulan Data	Tahapan Pengumpulan data yang tepat yaitu dengan mempertimbangkan penggunaannya berdasarkan jenis data dan sumbernya. Data yang objektif dan relevan dengan pokok permasalahan penelitian yang merupakan indikator keberhasilan suatu penelitian.
Pengelolaan Data	Tahapan ini menjelaskan tahap awal dalam data mining. Data terkumpul dan dinyatakan lengkap, data tersebut kemudian di seleksi atas dasar reliabilitas dan validitas. Pengelolaan data juga harus memperhatikan bentuk rumusan masalah dan tujuan penelitian.
Identifikasi Masalah	Tahap identifikasi masalah adalah mengelompokkan permasalahan yang di peroleh sesuai dengan tujuan penelitian.
Pemodelan	Tahapan ini Menjelaskan tentang pemilihan dan pemodelan metode yang digunakan dengan metode algoritma C4.5 pada penelitian. Penjelasan ini juga mengenai pengaturan untuk mendukung proses dan pemilihan nilai dari atribut – atribut serta tahapan proses pemodelan dengan algoritma C4.5 yang digunakan
Pengujian Model	Tahapan ini Menjelaskan mengenai evaluasi dan validitas hasil penerapan metode algoritma C4.5 pada penelitian yang dilakukan.
Evaluasi dan Validasi	Tahapan terakhir ini menjelaskan mengenai evaluasi dan validasi hasil penerapan metode algoritma C4.5 pada penelitian yang dilakukan.

3.1. Sumber Data

Data dalam penelitian ini yaitu Data primer yang diperoleh adalah sumber data yang langsung memberikan data kepada pengumpul data. Dalam

penelitian ini, sumber data primer diperoleh melalui wawancara dan observasi langsung ke konter feeyza cell. Berikut contoh data yang di peroleh di konter feeyza cell bisa lihat di tabel 2

Tabel 2. Data Pembelian

NO	Tanggal	Jenis Kartu Perdana	Qty	Harga	Jumlah Terjual	sis	Total	Status Penjualan
1	01-Sep-23	Telkomsel	250	Rp 8.000	3	247	Rp 24.000	Laris
2		Indosat	250	Rp 6.000	4	246	Rp 24.000	Laris
3		3	250	Rp 5.000	2	248	Rp 10.000	Laris
4		XL	250	Rp 7.000	1	249	Rp 7.000	Laris
5	02-Sep-23	Telkomsel	247	Rp 8.000	5	242	Rp 40.000	Laris
6		Indosat	246	Rp 6.000	3	243	Rp 18.000	Laris
7		3	248	Rp 5.000	1	247	Rp 5.000	Laris
8		XL	249	Rp 7.000	2	247	Rp 14.000	Laris
9	03-Sep-23	Telkomsel	242	Rp 8.000	4	238	Rp 32.000	Laris
10		Indosat	243	Rp 6.000	2	241	Rp 12.000	Laris
11		3	247	Rp 5.000	5	242	Rp 25.000	Laris
12		XL	247	Rp 7.000	1	246	Rp 7.000	Laris
...
...
361	30-Nov-23	Telkomsel	24	Rp 8.000	1	23	Rp 8.000	Laris
362		Indosat	64	Rp 6.000	0	64	Rp -	Tidak Laris
363		3	108	Rp 5.000	0	108	Rp -	Tidak Laris
364		XL	99	Rp 7.000	0	99	Rp -	Tidak Laris

3.2. Populasi dan Sampel

3.2.1. Populasi

Populasi terdiri dari obyek/subyek yang mempunyai sifat-sifat atau ciri-ciri tertentu yang ditetapkan oleh peneliti untuk dipelajari dan diambil kesimpulannya. Itu merupakan wilayah yang digeneralisasikan. Komentar di atas adalah bermanfaat. Hal ini untuk penulis menentukan

populasinya. Populasi yang digunakan dalam penelitian adalah Feyza cell counter.

3.2.2. Sampel

Ini adalah bagian dari populasi yang ingin diteliti oleh peneliti. Menurut Sugiyono (2011: 81), "Sampel adalah himpunan bagian dari jumlah dan ciri-ciri yang dimiliki suatu populasi." Karena sampel adalah bagian dari populasi yang ada, maka

pengambilan sampel melibatkan pertimbangan yang ada. menggunakan metode tertentu berdasarkan Dalam metode pengambilan sampel ini, penulis menggunakan teknik purposive sampling sampel yang ditargetkan adalah teknik penentuan sampel dari aspek tertentu. Berdasarkan pengertian di atas, maka dimungkinkan untuk memajukan penelitian. Oleh karena itu, penulis memutuskan untuk mengambil sampel penjualan kartu perdana di Sel Phaser. Pengambilan counter.

3.3. Teknik Pengumpulan Data

3.3.1. Wawancara

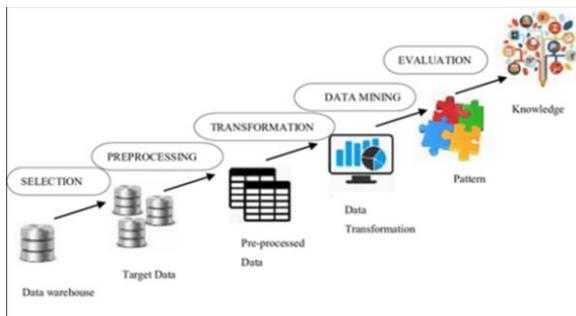
Wawancara adalah suatu metode pengumpulan data dimana responden diminta melakukan sesuatu dalam percakapan pribadi. Wawancara adalah bentuk pengumpulan data yang paling umum digunakan dalam penelitian kuantitatif.

3.3.2. Observasi

Observasi dilakukan dengan cara melihat dan mengamati permasalahan yang terjadi untuk memperoleh data yang diterima.

3.4. Teknik Analisis Data

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



Gambar 2. Tahapan Metode KDD

3.4.1. Selection (seleksi)

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

3.4.2. Preprocessing (pemilihan data)

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

3.4.3. Data transformation

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

3.4.4. Data mining

Data mining adalah proses mengeksplorasi dan menganalisis data dalam jumlah besar dengan menggunakan teknik dan metode tertentu dengan tujuan menemukan pola dan informasi menarik dalam data yang disimpan dalam jumlah besar. Teknik, metode, dan algoritma yang tepat sebenarnya bergantung pada tujuan dan proses penemuan pengetahuan di seluruh database (KDD). Fase ini merupakan inti dari fase Knowledge Discovery in Database (KDD) dan dilakukan untuk menganalisis data bersih.

3.4.5. Evaluation (evaluasi)

Penerjemah alur yang dihasilkan dari data mining, alur informasi yang dihasilkan dari proses perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Proses ini merupakan proses KDD yang mencakup pemeriksaan apakah alur atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada pada sebelumnya.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Hasil penelitian yang di tuangkan di dalam pembahasan ini yaitu memaparkan bagaimana menganalisis data set penjualan kartu perdana menggunakan Algoritma C4.5. Penelitian ini dilakukan melalui tahapan pengujian machine learning yaitu software RapidMiner studio.

4.2. Lokasi penelitian

Penelitian ini dilakukan di Konter FEEYZA CELL yang terletak pada Jl. Raya Laswi Kp.Pelangi Kec.Majalaya Kab.Bandung Jawa Barat 40382.

4.3. Data selection

Data yang digunakan dalam penelitian ini adalah data penjualan kartu perdana pada Bulan September - November 2023. Kemudian data tersebut diseleksi dan akan digunakan untuk diolah menggunakan aplikasi/tools Rapidminer. Tahapan seleksi dilakukan di Microsoft Excel dengan atribut yang digunakan diantaranya No, Tanggal, Jenis Kartu Perdana, Qty, Harga, Jumlah Terjual, sisa, Total, dan Status Penjualan. Hasil seleksi data menghasilkan 364 data.

Tabel 3. Data Penjualan

NO	Tanggal	Jenis Kartu Perdana	Qty	Harga	Jumlah Terjual	sis	Total	Status Penjualan
1	01-Sep-23	Telkomsel	250	Rp 8.000	3	247	Rp 24.000	Laris
2		Indosat	250	Rp 6.000	4	246	Rp 24.000	Laris
3		3	250	Rp 5.000	2	248	Rp 10.000	Laris
4		XL	250	Rp 7.000	1	249	Rp 7.000	Laris
5	02-Sep-23	Telkomsel	247	Rp 8.000	5	242	Rp 40.000	Laris
6		Indosat	246	Rp 6.000	3	243	Rp 18.000	Laris
7		3	248	Rp 5.000	1	247	Rp 5.000	Laris
8		XL	249	Rp 7.000	2	247	Rp 14.000	Laris
9	03-Sep-23	Telkomsel	242	Rp 8.000	4	238	Rp 32.000	Laris
10		Indosat	243	Rp 6.000	2	241	Rp 12.000	Laris
11		3	247	Rp 5.000	5	242	Rp 25.000	Laris
12		XL	247	Rp 7.000	1	246	Rp 7.000	Laris
...
...
361	30-Nov-23	Telkomsel	24	Rp 8.000	1	23	Rp 8.000	Laris
362		Indosat	64	Rp 6.000	0	64	Rp -	Tidak Laris
363		3	108	Rp 5.000	0	108	Rp -	Tidak Laris
364		XL	99	Rp 7.000	0	99	Rp -	Tidak Laris

4.4. Data preprocessing

Pada tahap ini dilakukan preprocessing data yaitu proses cleaning atau pembersihan data. Proses penghapusan data atribut yang tidak diperlukan seperti No, Pembeli dan Tanggal Transaksi. Sedangkan atribut yang akan digunakan adalah No, Tanggal, Jenis Kartu Perdana dan Qty.

Tabel 4. Data Preprocessing

NO	Tanggal	Jenis Kartu Perdana	Qty
1	01-Sep-23	Telkomsel	250
2		Indosat	250
3		3	250
4		XL	250
5	02-Sep-23	Telkomsel	247
6		Indosat	246
7		3	248
8	03-Sep-23	XL	249
9		Telkomsel	242
10		Indosat	243
11	04-Sep-23	3	247
12		XL	247
13		Telkomsel	238
14	05-Sep-23	Indosat	241
15		3	242
16		XL	246
17	30-Nov-23	Telkomsel	236
18		Indosat	239
19		3	240
20		XL	244
...
...
361	30-Nov-23	Telkomsel	24
362		Indosat	64
363		3	108
364		XL	99

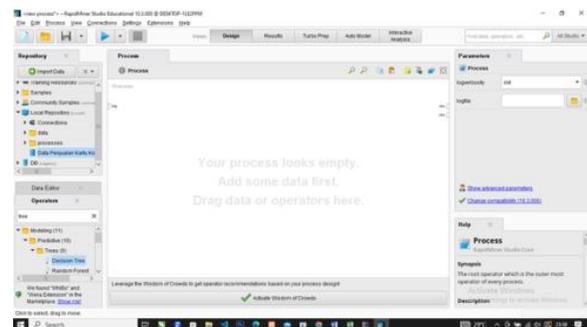
4.4.1. Data Mining

Teknik analisis data yang digunakan dalam penelitian ini adalah Perhitungan Data Menggunakan

Algoritma C4.5 dengan Rapidminer. Analisis yang dilakukan yaitu melakukan proses analisis dengan menggunakan tahapan Algoritma C4.5 sampai dengan membentuk akhir dari pohon keputusan dengan software rapidminer. Berikut pengaplikasian metode C4.5 dalam aplikasi/tools Rapidminer Diantaranya sebagai berikut:

4.4.2. Tampilan halaman utama Rapidminer

Pada gambar dibawah adalah tampilan halaman utama Rapidminer ketika dibuka



Gambar 3. Halaman Utama

RapidMiner adalah salah satu software untuk pengolahan data mining. Pekerjaan RapidMiner Text Mining berpusat pada analisis teks dan pola dari kumpulan data besar dan menggabungkannya dengan metode statistik, kecerdasan buatan, dan database. Tujuan dari analisis teks ini adalah untuk memperoleh informasi dengan kualitas terbaik dari teks yang diolah. RapidMiner menyediakan teknik penambangan data dan pembelajaran mesin seperti ETL (ekstrak, transformasi, muat), prapemrosesan data, visualisasi, pemodelan, dan evaluasi. Proses penambangan data ditulis menggunakan XML dan terdiri dari operator bersarang yang dibuat menggunakan GUI. Presentasi ditulis dalam bahasa pemrograman Java.

4.4.3. Proses input data dengan Format excel

Retrieve Data Penjualan Kartu Konter Feeyza Cell



Gambar 4. Retrive Data

Pada gambar 4 menjelaskan proses penginputan data. Data yang diinput bisa berbagai macam format seperti CSV, Excel dan masih banyak lainnya. Akan tetapi file yang akan diuji di aplikasi Rapidminer ini menggunakan data dengan format Excel yang sudah disiapkan.

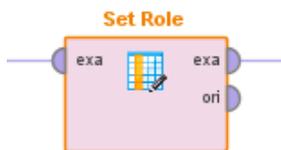
4.4.4. Decision Tree (Pohon Keputusan)



Gambar 5. Decision Tree

Pada gambar 5 Operator Decision Tree menghasilkan model pohon keputusan, yang dapat digunakan untuk klasifikasi dan regresi.

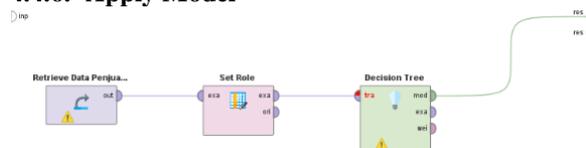
4.4.5. Set Role



Gambar 6. Set Role

Peran Atribut menjelaskan bagaimana Operator lain menangani Atribut ini. Peran default adalah reguler, peran lainnya diklasifikasikan sebagai khusus. Sebuah ContohSet dapat memiliki banyak Atribut khusus, namun setiap peran khusus hanya dapat muncul satu kali. Jika peran khusus ditetapkan ke lebih dari satu Atribut, semua peran akan diubah menjadi peran biasa kecuali Atribut terakhir. Berbagai jenis peran dijelaskan di bawah pada bagian parameter.

4.4.6. Apply Model



Gambar 7. Apply Model

Apply model merupakan tahapan akhir yang mana pada proses ini akan menerapkan model data mining C4.5 berdasarkan proses-proses sebelumnya mulai dari Read excel, Set Role dan Decision Tree. Setelah semua proses dilakukan, maka akan didapatkan hasil Pohon Keputusan dari data tersebut. Pada gambar 4 di bagian operator Retrieve data akan dimasukan seperti gambar 8

Row No.	Status Peng...	NO	Tanggal	Jenis Kartu ...	Qty	Harga	Jumlah Terj...	Sisa	Total	
1	Laris	1	Sep 1, 2023	Telkomsel	250	8000	3	247	24000	
2	Laris	2	?	Indosat	250	6000	4	246	24000	
3	Laris	3	?	?	3	250	5000	2	248	10000
4	Laris	4	?	XL	250	7000	1	249	7000	
5	Laris	5	Sep 2, 2023	Telkomsel	247	8000	5	242	40000	
6	Laris	6	?	Indosat	246	6000	3	243	18000	
7	Laris	7	?	?	3	248	5000	1	247	5000
8	Laris	8	?	XL	249	7000	2	247	14000	
9	Laris	9	Sep 3, 2023	Telkomsel	242	8000	4	238	32000	
10	Laris	10	?	Indosat	243	6000	2	241	12000	
11	Laris	11	?	?	3	247	5000	5	242	25000
12	Laris	12	?	XL	247	7000	1	246	7000	
13	Laris	13	Sep 4, 2023	Telkomsel	238	8000	2	236	16000	

Gambar 8. Memasukan Data di Retrive Data

4.5. Pembahasan

4.5.1. Penerapan Metode C4.5

Di bawah ini adalah penjelasan detail setiap langkah dalam membangun pohon keputusan dengan menggunakan algoritma C4.5:

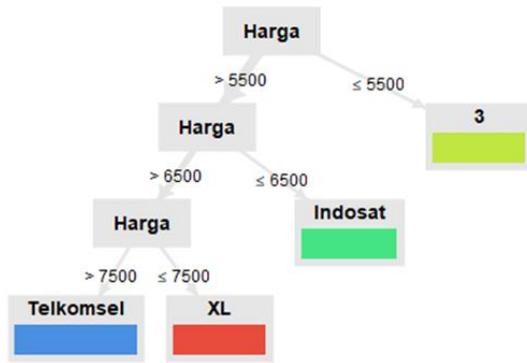
- Menghitung : jumlah kasus untuk keputusan Kurang Laris, jumlah kasus untuk keputusan Laris, Menghitung entropy untuk semua kasus, untuk sub kasus (berdasarkan atribut) dan menghitung gain untuk setiap atribut.

Hasil perhitungan untuk node awal dapat dilihat pada table 5

Tabel 5. Perhitungan Node Awal

Node		Jumlah Kasus	Status penjualan		Nilai	
			Laris	Kurang Laris	Entropy	Gain
		(S)	(S1)	(S2)		
1	Total	1000	706	294	0,016976999	
	Jenis Kartu Perdana					0,026693448
	Telkomsel	250	227	23	0,177448657	
	Indosat	250	186	64	0,076615303	
	3	250	142	108	-0,292929755	
	XL	250	151	99	-0,19815419	
	Harga					2,310770059
	<=5000	250	142	108	-0,292929755	
	>5000	750	564	186	0,087678385	
	Jumlah Terjual					0,811000649
	<= 3	276	276	38	0,393850304	
	>3	290	52	238	-2,245497869	

- Pengaplikasian Pohon Keputusan metode C4.5 dalam aplikasi/tools Rapidminer Diantaranya sebagai berikut :



Gambar 9. Pohon Keputusan di Hasilkan Rapid Miner

- c. Data Description Pohon keputusan metode C4.5 dalam aplikasi/tools Rapidminer Diantaranya sebagai berikut :

```

Tree
E > 5500
| E > 6500
| | E > 7500: Telkomsel {Telkomsel=91,
Indosat=0, 3=0, XL=0}
| | E ≤ 7500: XL {Telkomsel=0, Indosat=0,
3=0, XL=91}
| E ≤ 6500: Indosat {Telkomsel=0,
Indosat=91, 3=0, XL=0}
E ≤ 5500: 3 {Telkomsel=0, Indosat=0, 3=91,
XL=0}
    
```

5 KESIMPULAN DAN SARAN

Berdasarkan analisis menggunakan metode C4.5 terhadap pola pembelian konsumen terhadap produk operator telekomunikasi, kesimpulan yang dapat ditarik adalah sebagai berikut: Identifikasi pola pembelian: C4.5 mampu mengidentifikasi pola pembelian konsumen berdasarkan berbagai atribut seperti harga, kualitas layanan, paket promosi, cakupan jaringan, dan lainnya. Segmentasi pasar: Dengan menggunakan C4.5, kita dapat mengidentifikasi segmen pasar yang berbeda berdasarkan preferensi konsumen, memungkinkan operator telekomunikasi untuk menyesuaikan strategi pemasaran dan produk mereka. Prediksi perilaku konsumen: Metode C4.5 memungkinkan untuk membuat prediksi tentang perilaku pembelian konsumen di masa depan berdasarkan pola pembelian yang teramati, membantu operator telekomunikasi dalam merencanakan strategi penjualan dan promosi. Penyempurnaan strategi pemasaran: Dengan pemahaman yang lebih baik tentang pola pembelian konsumen, operator telekomunikasi dapat meningkatkan efektivitas kampanye pemasaran, menyesuaikan produk dan layanan mereka sesuai dengan kebutuhan dan preferensi pelanggan.

Hasil analisis C4.5 memberikan wawasan yang berharga bagi operator telekomunikasi untuk

meningkatkan strategi pemasaran, mengoptimalkan penjualan, dan meningkatkan kepuasan pelanggan.

DAFTAR PUSTAKA

- [1] I. Nawangsih and A. Asti Setyaningsih, "Penerapan Algoritma Naive Bayes Untuk Menentukan Klasifikasi Produk Terlaris Pada Penjualan Pulsa," *SIGMA –Jurnal Teknol. Pelita Bangsa*, vol. 10, no. 4, 2020.
- [2] G. Ramadhani and S. Widodo, "Pengaruh Citra Merek Terhadap Keputusan Pembelian Kartu Perdana Tri Pada Warga Kelurahan Pondok Karya Kota Tangerang Selatan," *J. Ilm. Mhs.*, vol. 1, no. 1, 2021, doi: 10.32493/jmw.v1i1.9517.
- [3] A. Tadjuk Gayoni, "PENGARUH HARGA TERHADAP KEPUTUSAN PEMBELIAN KONSUMEN KARTU SELULER TELKOMSEL," *J. Util.*, vol. 7, no. 1, 2022, doi: 10.22236/utilitas.v7i1.8299.
- [4] D. Pingsimoni, I. Samsie, N. Aini, P. Studi Rekayasa Perangkat Lunak, U. Dipa, and M. Jalan Perintis Kemerdekaan, "Penerapan Data Mining Menggunakan Algoritma C.45 Dalam Menentukan Pola Pembelian Pakaian Syar'i Yang terjual," vol. 1, 2022.
- [5] M. Sulaiman and T. G. Laksana, "Perbandingan algoritma c4.5 dengan c4.5 berbasis bagging dalam menganalisa pelanggan pulsa elektronik," *Semin. Nas. Sist. Inf. Indones.*, no. January 2015, 2018.
- [6] S. M. Putri and S. A. Arnomo, "Penerapan Algoritma C4.5 Untuk Prediksi Kualitas Pelayanan Terhadap Kepuasan Konsumen (Studi Kasus: Hinet Batam)," *J. Inf. Syst. Res.*, vol. 1, no. 2, 2020.
- [7] Y. Yudianta, A. Yulia Agustina, and dan Nur Khofifah, "Prediksi Customer Churn Menggunakan Metode CRISP-DM Pada Industri Telekomunikasi Sebagai Implementasi Mempertahankan Pelanggan," *Indones. J. Islam. Econ. Bus.*, vol. 8, no. 1, 2023.
- [8] A. Shiddiq, R. K. Niswatin, and I. N. Farida, "Analisa Kepuasan Konsumen Menggunakan Klasifikasi Decision Tree," *Gener. J.*, vol. 2, no. 1, 2018.
- [9] D. D. N. Kustianti, "Pengaruh Citra Merek dan Harga Terhadap Keputusan Pembelian Ulang Kartu Seluler Telkomsel," *Psikoborneo J. Ilm. Psikol.*, vol. 7, no. 1, 2019, doi: 10.30872/psikoborneo.v7i1.4709.
- [10] S. Adrianto, N. Khasanah, and D. Wahyuni, "Implementasi Data Mining pada Penjualan Kartu Perdana Internet di Purnama Ponsel Menggunakan Metode Algoritma Apriori," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 5, no. 2, 2020, doi: 10.14421/jiska.2020.52-03.