



Computer Based Information System Journal

ISSN (Print): 2337-8794 | E- ISSN : 2621-5292
web jurnal : <http://ejournal.upbatam.ac.id/index.php/cbis>



ANALISIS POTENSI PENINGKATAN PENJUALAN TOKO MITRA LAZADA AZGALERY DENGAN ALGORITMA C4.5

Rico Fransisco, Erlin Elisa

Universitas Putera Batam, Indonesia.

INFORMASI ARTIKEL

Diterima Redaksi: Januari 2024
Diterbitkan Online: Maret 2024

KATA KUNCI

Potensi, Analisis, Keputusan,
Algoritma C4.5

KORESPONDENSI

E-mail: rico_fransisco@hotmail.com

A B S T R A C T

The potential for sales in online marketplaces is increasing over time due to the development and distribution of telecommunications signals and equipment. This event is an opportunity that can be exploited by partner shops selling on the marketplace, one of which is Azgalery Lazada partners. Azgalery has been around for 4 years with a total of almost 700 products in various categories ranging from women's fashion, men's fashion, home decoration to other household equipment. This research aims to help partner shops get a decision tree that can be used for making crucial decisions such as preparing the latest products and price research. Apart from that, this also increases the customer trust in the azgalery store itself so that they are more trustworthy in the eyes of customers. The method used is an implementation using data mining, especially classification using the C4.5 algorithm. This method is used to analyze sales potential with the help of existing data to support research to produce valid results.

I. Latar Belakang

Penggunaan teknologi setiap harinya membuat tumpukan data semakin besar dan kompleks setiap harinya. Tumpukan data yang besar ini mencakup berbagai jenis informasi, mulai dari data transaksi bisnis, interaksi di media sosial, informasi kesehatan, hingga data lingkungan. Data tersebut tentunya dapat

digunakan dalam berbagai bidang, seperti dapat dimanfaatkan kembali untuk analisis proses usaha di waktu yang akan datang.

Salah satunya data yang telah ada tersebut dapat digunakan untuk menganalisis transaksi penjualan. Berbicara mengenai usaha bisnis dibidang E-commerce saat ini tentu sangat menjanjikan, Karena di era digitalisasi

penggunaan teknologi sangat berpengaruh terhadap kelangsungan usaha. Jika dilihat dari data estimasi nilai transaksi dari penjualan di E-commerce Indonesia setiap tahunnya pasti selalu mengalami lonjakan naik yang signifikan. Setiap harinya pasti selalu terjadi transaksi. Hal ini dapat dibuktikan dari gambar dibawah yang menjelaskan estimasi nilai transaksi E-commerce dari tahun 2021 hingga 2025. E-commerce (perdagangan elektronik) adalah salah satu bidang yang telah mengalami pertumbuhan luar biasa berkat kemajuan teknologi. E-commerce telah mengubah cara orang berbelanja dan berbisnis dengan memanfaatkan sebuah platform yang berbasis online. E-commerce merupakan salah satu pilihan terbaik bagi para pelaku usaha karena bisnis dapat dibantu beroperasi selama 24 jam tanpa batas waktu dalam menerima pesanan. Pertumbuhan ini juga turut berkontribusi pada tumpanyan data yang semakin besar karena setiap transaksi, interaksi, dan kegiatan lainnya di platform E-commerce menghasilkan data yang signifikan.

Salah satunya yang sedang populer akhir – akhir ini adalah belanja online di media E-Commerce yaitu marketplace. Lazada merupakan salah satu layanan belanja online yang terpercaya di Indonesia dengan jumlah kunjungan 82,5 juta orang pada kuartal pertama tahun 2023. Salah satu toko yang berjualan di platform Lazada adalah azgalery. Data transaksi di toko ini setiap bulannya akan diunduh dari marketplace kemudian hanya disimpan dalam server kantor saja. Data tersebut seharusnya diolah dan digunakan untuk mengambil keputusan krusial dengan pertimbangan yang baik di masa depan, tetapi faktanya hanya dibiarkan menumpuk di dalam server setiap bulannya. Jika data tersebut diolah kemudian menghasilkan sebuah data baru yang dapat dipakai untuk menentukan keputusan yang tepat, Sudah dapat dipastikan toko azgalery akan mengalami peningkatan penjualan yang baik setiap bulannya karena satu keputusan yang tepat diambil. Maka dari itu data tersebut sebaiknya diolah menggunakan metode yang tepat dan digunakan untuk mengambil keputusan yang tepat juga. Permasalahan pada penelitian sebelumnya dapat diselesaikan dengan salah satu metode data mining.

Metode yang dipakai untuk membantu mengolah data tersebut adalah Algoritma C4.5 Metode ini digunakan karena menghasilkan suatu

pohon keputusan yang menghasilkan sekelompok aturan-aturan yang bisa diimplementasikan untuk kepentingan toko tersebut kedepannya.

II. Kajian Literatur

A. Teori umum

Knowledge discovery in database ini sering dikaitkan dengan data mining karena memiliki arti yang sama yaitu pengumpulan data dalam set data yang berukuran besar kemudian diolah dengan sebuah proses menjadi sebuah informasi yang berharga [1]. Data yang berjumlah besar sangat berpotensi datanya menghilang dan tidak konsisten kemudian hal tersebut akan sangat berpengaruh pada proses data mining [2]. Data mining merupakan proses pengumpulan data kemudian dianalisis beberapa bagian dari kumpulan data besar yang sudah ada setelah itu dikembangkan menjadi sebuah data baru yang bermanfaat bagi penggunaanya [3]. Teknik ini sangat menguntungkan dipakai untuk menyelidiki sebuah informasi dan mengembangkannya menjadi data yang penting dan bermanfaat. Informasi yang ditambang dari database besar menurut banyak peneliti dianggap sebagai topik penelitian utama dalam sistem database yang penting dan memiliki peluang pendapatan yang besar [4]. Sekarang digunakan untuk berbagai kumpulan data untuk mempelajari pola tersembunyi dan membuat prediksi dan/atau deskripsi yang tepat. Penambangan data adalah kumpulan metode yang digunakan untuk mengekstrak pengetahuan tersembunyi seperti pola, hubungan, atau aturan dari kumpulan data yang besar [5].

Data mining juga menjadi salah satu metode penyelidikan yang berkembang sangat cepat dalam dunia teknologi informasi dikarenakan proses pengumpulan informasi yang diolah dengan baik untuk menunjang keputusan yang krusial di masa yang akan datang membuatnya menjadi metode yang sangat dipertimbangkan untuk digunakan. Klasifikasi data merupakan langkah untuk menemukan bagian sama dalam basis data dan memisahkannya ke dalam kelas yang berbeda menurut model yang ditetapkan [6].

dikarenakan mampu menghasilkan decision tree yang diimplementasikan dan diambil sebuah keputusan di masa yang akan datang [7]. Hal ini merupakan proses klasifikasi data dengan sistem kerja penyusunan langkah yang sesuai untuk

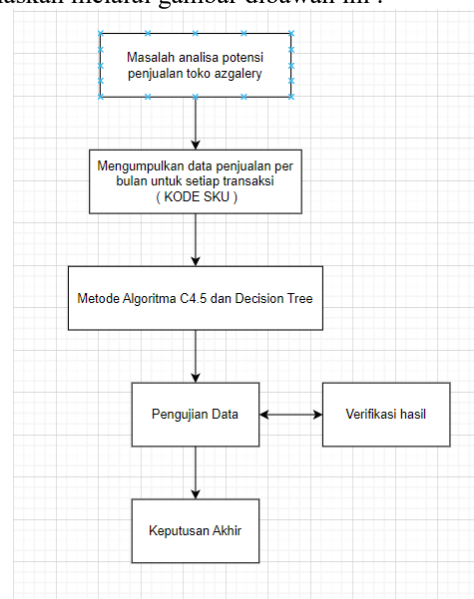
mengatasi sebuah masalah [8]. Dengan cara yang sama seperti Dichotomizer 3 (ID3) berulang, itu menggunakan konsep entropi informasi dari serangkaian pelatihan data untuk membangun pohon keputusan. Anda dapat melakukannya dengan menghitung nilai gain dari setiap atribut, dengan akar pertama menghasilkan nilai gain yang paling besar. Decision tree adalah alat pendukung keputusan yang direpresentasikan dalam bentuk struktur pohon yang memiliki node, branch dan leaf. Alat yang mendukung dengan struktur seperti pohon yang mendefinisikan berbagai probabilitas atau hasil dari sebuah pernyataan. Sebuah pohon keputusan merupakan rangka sederhana yang dihasilkan melalui salah satu metode yaitu algoritma C4.5 pada setiap node non-terminal adalah tes dan dianggap menjadi item data [7]. Mereka terdiri dari subbagian yang menunjukkan proses pengambilan keputusan yang dapat menghasilkan hasil yang menguntungkan. Struktur flowchart memiliki node internal yang mewakili tes atau fitur pada setiap tahap proses. Setiap cabang menunjukkan hasil atribut, dan jalur dari daun ke akar menunjukkan aturan klasifikasi. Karena alat ini dapat menangani masalah penyesuaian data seperti regresi dan klasifikasi, itu juga berguna untuk menyesuaikan hubungan non-linier [9]. Rapidminer merupakan salah satu perangkat lunak yang dikembangkan untuk membantu proses perhitungan dalam data mining. Rapidminer merupakan bagian dari perangkat lunak untuk mengolah data dan membuat bentuk dari set data yang besar menggunakan salah satu prinsip dalam data mining dan penggabungan metode lainnya. Aplikasi ini sangat membantu dalam perhitungan data set yang besar kemudian menghasilkan hasil dalam waktu yang lebih singkat [10].

B. Teori khusus

Penjualan merupakan proses yang melibatkan penawaran dalam bentuk produk ataupun jasa dari penjual kepada pembeli dengan nilai tukar tertentu [11]. dalam kasus marketplace ini adalah penjualan sebuah produk yang dipasarkan online kemudian dapat dilihat oleh calon pelanggan dengan cara mencari melalui kata kunci yang sesuai. Transaksi yang terjadi Ketika pelanggan membeli barang dengan berbagai macam jenis pembayaran, hal itulah yang dikatakan penjualan. Produk merupakan barang atau jasa yang dijual kepada pelanggan. Produk

bisa dikatakan sebagai nilai tukar yang dimiliki oleh penjual yang ditawarkan kepada pembeli. Produk di marketplace banyak jenis dan kategorinya dan cakupannya luas dari berbagai daerah [12]. Produk yang paling banyak menarik minat di marketplace adalah fashion, dekorasi kemudian peralatan rumah tangga karena hal tersebut dinilai praktis dan lebih memiliki potensi penjualan yang baik daripada produk berkategori elektronik, otomotif, Kesehatan dikarenakan cenderung ditawarkan dengan harga yang mahal dan butuh tingkat kepercayaan dari calon pembeli yang tinggi [13].

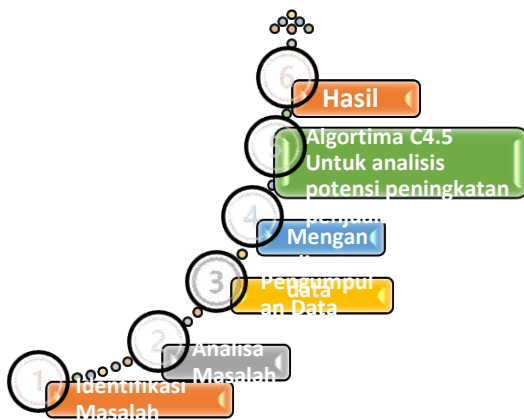
Kerangka berpikir dari penelitian ini dapat dijelaskan melalui gambar dibawah ini :



Gambar 1. Kerangka Pemikiran

III. Metodologi

Berikut adalah gambaran desain penelitian untuk menjelaskan analisis potensi peningkatan penjualan pada toko mitra lazada azgalery.



Gambar 2. Desain Penelitian

Dari langkah desain penelitian yang akan di laksanakan sesuai dengan alur gambar diatas maka kegiatan penelitian ini dapat diklasifikasikan seperti dibawah ini.

1. Identifikasi masalah

Mengidentifikasi setiap permasalahan yang terjadi pada objek penelitian dimana pada kesempatan ini peneliti akan melakukan survei awal di lokasi penelitian dengan menemukan masalah yang ada pada objek penelitian, tentang bagaimana selama ini data disimpan dan dikelola kemudian penulis mencari sumber masalah yang sedang terjadi kemudian urgensinya untuk dilakukan penelitian dan dicarikan solusi.

2. Analisa masalah

Identifikasi permasalahan yang ditemukan kegiatan selanjutnya yang akan dilaksanakan analisis permasalahan, dimana pada tahap ini penulis akan melakukan uraian permasalahan yang terjadi kemudian melihat kesempatan terhadap permasalahan yang utama di selesaikan terkait dengan data transaksi penjualan yang menjadi urgensi pada penelitian ini.

3. Pengumpulan Data

Proses pengumpulan data dimulai dari mengumpulkan data transaksi penjualan selama

<http://ejournal.upbatam.ac.id/index.php/cbis>

ini yang telah dilakukan oleh objek penelitian, dalam hal ini penulis meminta izin kepada objek penelitian untuk memberikan data-data yang dibutuhkan guna melanjutkan kegiatan pada langkah selanjutnya.

4. Meninjau dan menganalisa data

Setelah data diperoleh dari objek penelitian, data-data tersebut akan ditinjau dan di analisis agar data-data yang di berikan dapat di sesuaikan dengan yang akan di analisis dalam mencari informasi atau pengetahuan baru terhadap potensi tingkat penjualan selama ini.

5 Menerapkan Algoritma C4.5

Pada bagian ini penulis mulai mengolah data yang telah dianalisis kemudian menerapkan salah satu teknik datamining algoritma C4.5

6. Hasil

Penulis akan menguji keberhasilan hasil yang di peroleh dari analisis algoritma C4.5 dengan memanfaatkan salah satu software datamining *open source* rapidminer.

IV. Pembahasan

A. Perhitungan Manual

Langkah awal dalam perhitungan manual adalah mengkalkulasikan nilai *entropy* setiap atribut kemudian menentukan nilai *gain* tertinggi dari atribut yang ada untuk diambil sebagai akar dari pohon keputusan.

Entropy (Total)

$$\begin{aligned}
 Entropy(total) &= \left(-\frac{73}{100} * \log_2 \left(\frac{73}{100} \right) \right) \\
 &+ \left(-\frac{27}{100} * \log_2 \left(\frac{27}{100} \right) \right) \\
 &= 0.8414
 \end{aligned}$$

a. Atribut Kualitas

– Entropy BK

$$= \left(-\frac{65}{80} * \log_2 \left(\frac{65}{80} \right) \right) + \left(-\frac{15}{80} * \log_2 \left(\frac{15}{80} \right) \right) = 0.6962$$

– Entropy CK

$$= \left(-\frac{8}{20} * \log_2 \left(\frac{8}{20} \right) \right) + \left(-\frac{12}{20} * \log_2 \left(\frac{12}{20} \right) \right) = 0.9709$$

Sedangkan untuk nilai gain kualitas sebagai berikut.

$$= 0.8414 - \left(\left(\frac{80}{100} * 0.6962 \right) + \left(\frac{20}{100} * 0.9709 \right) \right) = 0.0903$$

b. Atribut Volume Pembelian

– Entropy BA

$$= \left(-\frac{25}{25} * \log_2 \left(\frac{25}{25} \right) \right) + \left(-\frac{0}{25} * \log_2 \left(\frac{0}{25} \right) \right) = 0$$

– Entropy SD

$$= \left(-\frac{48}{75} * \log_2 \left(\frac{48}{75} \right) \right) + \left(-\frac{27}{75} * \log_2 \left(\frac{27}{75} \right) \right) = 0.9426$$

Sedangkan untuk nilai gain volume pembelian adalah sebagai berikut, seperti

dibawah ini: $= 0.8414 - \left(\left(\frac{25}{100} * 0 \right) + \left(\frac{75}{100} * 0.9426 \right) \right) = 0.1344$

c. Atribut Harga

– Entropy MR

$$= \left(-\frac{23}{30} * \log_2 \left(\frac{23}{30} \right) \right) + \left(-\frac{7}{30} * \log_2 \left(\frac{7}{30} \right) \right) = 0.7837$$

– Entropy SE

$$= \left(-\frac{32}{48} * \log_2 \left(\frac{32}{48} \right) \right) + \left(-\frac{16}{48} * \log_2 \left(\frac{16}{48} \right) \right) = 0.9182$$

– Entropy MA

$$= \left(-\frac{18}{22} * \log_2 \left(\frac{18}{22} \right) \right) + \left(-\frac{4}{22} * \log_2 \left(\frac{4}{22} \right) \right) = 0.6840$$

Sedangkan untuk nilai gain harga adalah sebagai berikut, seperti dibawah ini:

$$= 0.8414 - \left(\left(\frac{30}{100} * 0.7837 \right) + \left(\frac{48}{100} * 0.9182 \right) + \left(\frac{22}{100} * 0.6840 \right) \right) = 0.0150$$

d. Atribut Diskon

– Entropy DK

$$= \left(-\frac{26}{30} * \log_2 \left(\frac{26}{30} \right) \right) \\ + \left(-\frac{4}{30} * \log_2 \left(\frac{4}{30} \right) \right) \\ = 0.5665$$

– Entropy TD

$$= \left(-\frac{47}{70} * \log_2 \left(\frac{47}{70} \right) \right) \\ + \left(-\frac{23}{70} * \log_2 \left(\frac{23}{70} \right) \right) \\ = 0.9134$$

Sedangkan untuk nilai *gain* diskon adalah sebagai berikut, seperti dibawah ini:

$$= 0.8414 - \left(\left(\frac{30}{100} * 0.5665 \right) + \left(\frac{70}{100} * 0.9134 \right) \right) = 0.0320$$

e. Atribut Delivery

– Entropy LB

$$= \left(-\frac{28}{47} * \log_2 \left(\frac{28}{47} \right) \right) \\ + \left(-\frac{19}{47} * \log_2 \left(\frac{19}{47} \right) \right) \\ = 0.9733$$

– Entropy CP

$$= \left(-\frac{45}{53} * \log_2 \left(\frac{45}{53} \right) \right) \\ + \left(-\frac{8}{53} * \log_2 \left(\frac{8}{53} \right) \right) \\ = 0.6121$$

Sedangkan untuk nilai *gain delivery* adalah sebagai berikut, seperti dibawah ini:

$$= 0.8414 - \left(\left(\frac{47}{100} * 0.9733 \right) + \left(\frac{53}{100} * 0.6121 \right) \right) \\ = 0.0595$$

Berikut merupakan tabel yang sudah dimasukkan hasil hitung dari semua *entropy* dan *gain*.

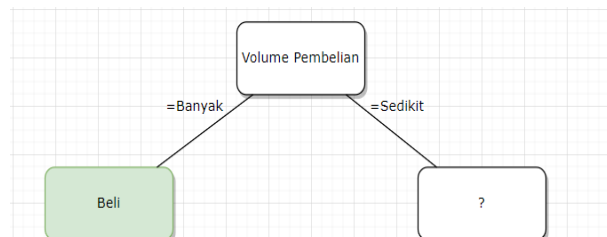
Tabel 1. Perhitungan Node 1 (manual)

	Keputusan		Beli	Tidak Beli	Entropy	Gain
Total		100	73	27	0.841464636	
Kualitas						
Baik	BK	80	65	15	0.69621226	0.090304709
Cukup	CK	20	8	12	0.970950594	
		100				
Volume Pembelian						
Banyak	BA	25	25	0	0	0.134452244
Sedikit	SD	75	48	27	0.942683189	

			100				
Harga							
Murah	MR	30	23	7	0.783776947	0.015061096	
Sedang	SE	48	32	16	0.918295834		
Mahal	MA	22	18	4	0.684038436		
			100				
Diskon							
Diskon	DK	30	26	4	0.566509507	0.032089683	
Tidak ada Diskon	TD	70	47	23	0.913460145		
			100				
Delivery							
Lambat	LB	47	28	19	0.973385435	0.059509534	
Cepat	CP	53	45	8	0.612196127		
			100				

Sumber: (Penulis, 2023)

Dari hasil perhitungan yang ada pada table diatas terlihat bawa gain terbesar ada pada volume pembelian. Sehingga menghasilkan Volume sebagai *node* akar pertama dimana volume pembelian yang banyak sudah dipastikan pelanggan membeli karena nilai entropy sama dengan 0. Sehingga menghasilkan pohon keputusan seperti gambar dibawah ini.



Gambar 3. Pohon Keputusan
Node 1

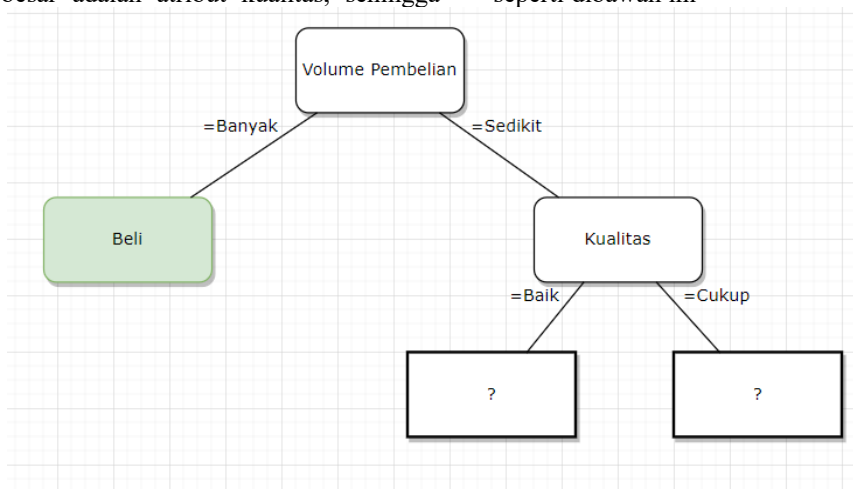
Tabel 2. Perhitungan Node 2 (manual)

Keputusan			Beli	Tidak Beli	Entropy	Gain
Total			75	48	27	0.942683
Kualitas						
Baik	BK	60	45	15	0.811278	0.149275
Cukup	CK	15	3	12	0.721928	
		75				
Harga						
Murah	MR	30	23	7	0.783777	0.063204

Sedang	SE	40	24	16	0.970951	
Mahal	MA	5	1	4	0.721928	
		75				
Diskon						
Diskon	DK	13	9	4	0.890492	
Tidak ada Diskon	TD	62	39	23	0.951412	0.001831
		75				
Delivery						
Lambat	LB	35	16	19	0.994694	
Cepat	CP	40	32	8	0.721928	0.093464
		75				

Sumber: (Penulis, 2023)

Dari perhitungan diatas yang memiliki *gain* terbesar adalah atribut kualitas, sehingga volume pembelian menjadi akar *node* kedua seperti dibawah ini



Gambar 4. Pohon Keputusan
Node 2

Tabel 3. Perhitungan Node 3 Kualitas cukup

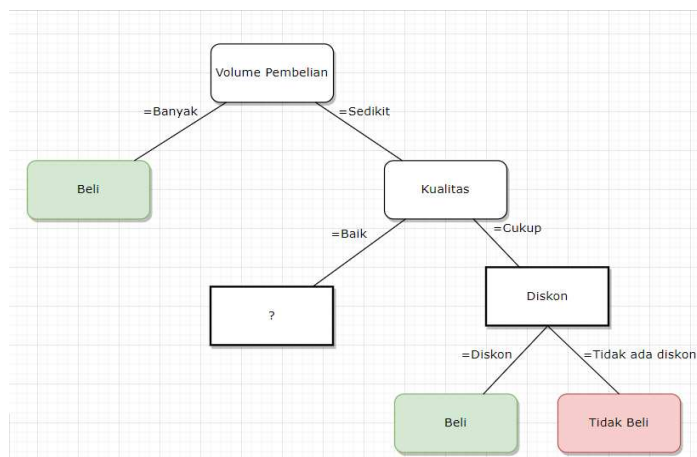
	Keputusan		Beli	Tidak Beli	Entropy	Gain
Total		15	3	12	0.721928	
Kualitas						
Baik	BK	0	0	0	0	0

Cukup	CK	15	3	12	0.721928	
		15				
Harga						
Murah	MR	8	1	7	0.543564	
Sedang	SE	6	1	5	0.650022	0.172018
Mahal	MA	1	1	0	0	
		15				
Diskon						
Diskon	DK	3	3	0	0	
Tidak ada Diskon	TD	12	0	12	0	0.721928
Delivery						
Lambat	LB	5	1	4	0.721928	
Cepat	CP	10	2	8	0.721928	0

Sumber: (Penulis, 2023)

Dari perhitungan *node* kualitas cukup didapati diskon merupakan nilai *gain* tertinggi.

Sehingga akar pohon keputusan yang dihasilkan seperti dibawah ini.



Gambar 5. Pohon Keputusan
Node 3 Baik

Tabel 4. Perhitungan Node 3 Kualitas baik

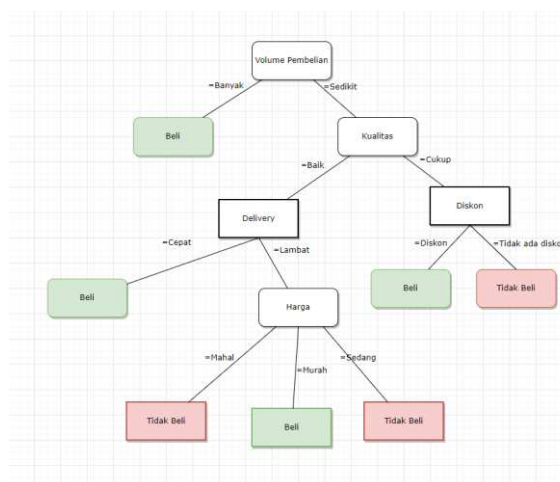
Keputusan		Beli	Tidak Beli	Entropy	Gain
Total	60	45	15	0.811278	
Harga					

Murah	MR	22	22	0	0	
Sedang	SE	34	23	11	0.908178	0.296644
Mahal	MA	4	0	4	0	
		60				
Delivery						
Lambat	LB	30	15	15	1	
Cepat	CP	30	30	0	0	
		60				

0.311278

Sumber: (Penulis, 2023)

Dari perhitungan diatas didapati nilai *gain* tertinggi adalah delivery sehingga menjadi node 3 dari akar pohon keputusan sehingga otomatis atribut harga menjadi node terakhir dalam akar keputusan dengan entropy yang sudah pasti membuat atribut harga tidak perlu dihitung kembali karena sudah ada keputusannya.



Gambar 6. Hasil pohon keputusan (*decision tree*)

B. Perhitungan Algoritma C4.5 Menggunakan RapidMiner

Setelah melakukan perhitungan manual, peneliti menggunakan aplikasi Rapidminer untuk menghitung apakah hasil yang didapatkan sudah valid atau tidak. Dibawah ini merupakan langkah – langkah perhitungannya.

1. Excel yang berisi 100 sampel data

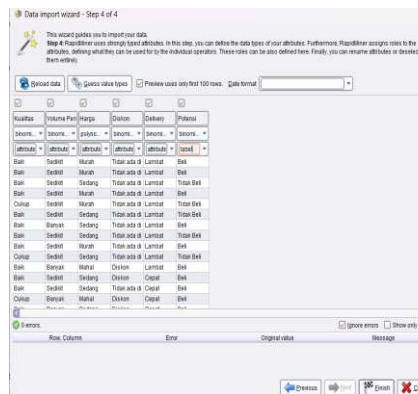
<http://ejournal.upbatam.ac.id/index.php/cbis>

yang akan diuji lengkap dengan atributnya yang sudah melalui random sampling dan mewakili data 3 tahun terakhir tersebut disimpan dengan nama *datatransaksi.xls* seperti dibawah ini.

	A	B	C	D	E	F	G	H
1	Kualitas	Volume Pembelian	Harga	Diskon	Delivery	Potensi		
2	Baik	Sedikit	Murah	Tidak ada diskon	Lambat	Beli		
3	Baik	Sedikit	Murah	Tidak ada diskon	Lambat	Beli		
4	Baik	Sedikit	Sedang	Tidak ada diskon	Lambat	Tidak Beli		
5	Baik	Sedikit	Murah	Tidak ada diskon	Lambat	Beli		
6	Cukup	Sedikit	Murah	Tidak ada diskon	Lambat	Tidak Beli		
7	Baik	Sedikit	Sedang	Tidak ada diskon	Lambat	Tidak Beli		
8	Baik	Banyak	Sedang	Tidak ada diskon	Lambat	Beli		
9	Baik	Sedikit	Sedang	Tidak ada diskon	Lambat	Tidak Beli		
10	Baik	Sedikit	Murah	Tidak ada diskon	Lambat	Beli		
11	Cukup	Sedikit	Sedang	Tidak ada diskon	Lambat	Tidak Beli		
12	Baik	Banyak	Mahal	Diskon	Lambat	Beli		
13	Baik	Sedikit	Sedang	Diskon	Cepat	Beli		
14	Baik	Sedikit	Sedang	Tidak ada diskon	Cepat	Beli		
15	Cukup	Banyak	Mahal	Diskon	Cepat	Beli		
16	Baik	Banyak	Sedang	Diskon	Cepat	Beli		
17	Baik	Sedikit	Murah	Tidak ada diskon	Cepat	Beli		
18	Cukup	Sedikit	Sedang	Tidak ada diskon	Cepat	Tidak Beli		
19	Baik	Sedikit	Murah	Tidak ada diskon	Cepat	Beli		
20	Cukup	Banyak	Sedang	Diskon	Cepat	Beli		
21	Cukup	Sedikit	Murah	Tidak ada diskon	Cepat	Tidak Beli		
22	Baik	Banyak	Sedang	Diskon	Cepat	Beli		
23	Cukup	Sedikit	Murah	Diskon	Cepat	Beli		
24	Baik	Banyak	Mahal	Diskon	Cepat	Beli		
25	Cukup	Sedikit	Sedang	Tidak ada diskon	Cepat	Tidak Beli		
26	Baik	Sedikit	Sedang	Tidak ada diskon	Cepat	Beli		
27	Baik	Sedikit	Murah	Tidak ada diskon	Cepat	Beli		
28	Baik	Banyak	Mahal	Diskon	Cepat	Beli		
29	Cukup	Sedikit	Sedang	Diskon	Cepat	Beli		

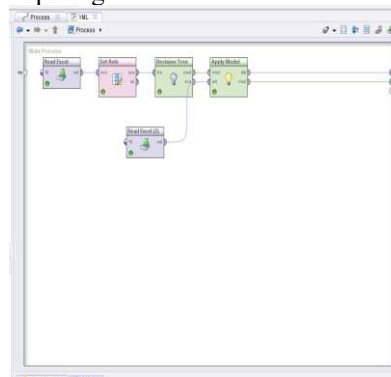
Gambar 7. *datatransaksi.xls*

2. Selanjutnya buka aplikasi Rapidminer 5 kemudian new process ,import data dan read excel. Masukkan excel yang sudah disiapkan yaitu *data-jadi.xls*. Pada bagian data import wizard , Atribut potensi harus dijadikan label karena adalah keputusan dalam data ini.



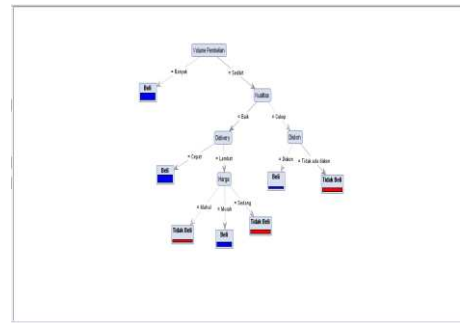
Gambar 8. Merubah atribut potensi menjadi label

3. Pada kolom search cari *set role* , setelah itu jadikan atribut potensi pada data sebagai label. Kemudian search decision tree untuk membuat pohon keputusan dan apply model. Hubungkan *output read excel* ke *set role* kemudian *transfer* ke decision tree dan yang terakhir *apply model*. Seperti gambar dibawah ini.



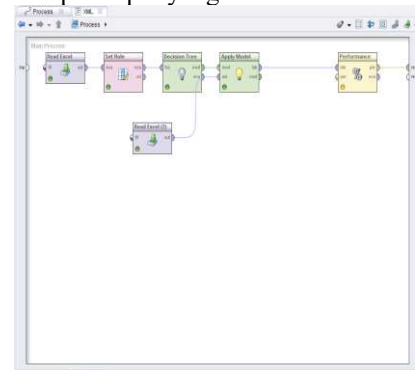
Gambar 9. Decision tree C4.5 dengan apply model

4. Lakukan *run* hingga didapati pohon keputusan seperti dibawah ini



Gambar 10. Pohon keputusan C4.5 dengan apply model

5. Pastikan hasil klasifikasi yang sudah dikeluarkan tadi dengan *performance* untuk menghitung berapa output yang benar dan salah.



Gambar 11. Menghubungkan hasil olah data dengan peformance

[illegible]

Gambar 12. Hasil peformance

Hasil performance decision tree pada gambar diatas menjelaskan bahwa peforma hasil data yang diolah sudah 100 % akurat tanpa kesalahan sehingga pengujian tersebut dinilai valid.

Setelah melewati proses perhitungan manual dan aplikasi rapidminer didapatkan pohon keputusan yang sama sehingga pohon keputusan yang dihasilkan melalui kedua metode tersebut menghasilkan aturan sebagai berikut.

1. Jika volume pembelian banyak , pelanggan sudah pasti membeli tanpa mempertimbangkan atribut lainnya.
2. Jika volume pembelian sedikit, maka pelanggan melihat atribut kualitas.
3. Jika kualitas cukup, maka pelanggan melihat apakah ada diskon yang diberikan oleh toko.
4. Jika diskon , maka pelanggan membeli barang tersebut.
5. Jika tidak ada diskon , maka pelanggan tidak membeli barang tersebut.
6. Jika kualitas baik, maka pelanggan melihat atribut delivery untuk memastikan apakah membeli atau tidak.
7. Jika delivery cepat, pelanggan akan langsung memutuskan membeli.
8. Jika delivery lambat, maka pelanggan mempertimbangkan atribut harga.
9. Jika harga mahal , maka pelanggan tidak membeli barang.
10. Jika harga sedang , pelanggan juga tidak membeli barang.
11. Jika harga murah , pelanggan sudah pasti membeli barang tersebut.

V. Kesimpulan

Berdasarkan penelitian yang sudah dilakukan maka diambil kesimpulan sebagai berikut :

a. Algoritma C4.5 memberikan pengetahuan tambahan terkait karakteristik pembeli dan faktor apa saja yang berpotensi untuk meningkatkan penjualan. Melalui implementasi strategi pemasaran yang disesuaikan dengan hasil decision tree untuk membantu toko menarik perhatian lebih dari pelanggan dan meningkatkan kinerja penjualan.

b. Decision tree yang dihasilkan dari RapidMiner memberikan pandangan jelas melalui aturan – aturan yang tercipta. Hal ini berguna sebagai panduan dari langkah yang nantinya bisa diambil toko untuk meningkatkan potensi penjualan.

Ucapan Terima Kasih

Peneliti ingin mengucapkan terimakasih kepada semua pihak yang telah membantu dalam proses penelitian dan jurnal dari awal hingga akhirnya selesai. Terutama kepada Ibu Erlin Elisa, S.Kom., M.Kom. selaku dosen pembimbing yang sudah membantu saya dalam keseluruhan proses. Tidak lupa juga peneliti berterimakasih kepada pihak toko azgalery yang sudah mengizinkan peneliti melakukan penelitian dan pengumpulan data dan kedua orang tua yang telah memberikan semangat pada peneliti dalam keseluruhan proses yang dijalani.

Daftar Pustaka

- [1] Yusuf Maulana, Riki Winanjaya, and Fitri Rizki, “Penerapan Data Mining dengan Algoritma C.45 Dalam Memprediksi Penjualan Tempe,” *Bulletin of Computer Science Research*, vol. 2, no. 2, pp. 53–58, Apr. 2022, doi: 10.47065/bulletincsr.v2i2.163.
- [2] S. A. & B. W. S. Alasadi, “Review of Data Preprocessing Techniques,” 2017.
- [3] C. Clifton, “Data Mining.” [Online]. Available: https://en.wikipedia.org/wiki/Data_mining
- [4] J. H. and P. S. Y. Ming-Syan Chen, “Data mining: an overview from a database perspective,” *IEEE Trans Knowl Data Eng*, vol. 8, pp. 866–883, 2022.
- [5] I. D. Mienye, Y. Sun, and Z. Wang, “Prediction performance of improved decision tree-based algorithms: A review,” in *Procedia Manufacturing*, Elsevier B.V., 2019, pp. 698–703. doi: 10.1016/j.promfg.2019.06.011.
- [6] N. Azwanti, “ANALISA ALGORITMA C4.5 UNTUK MEMPREDIKSI PENJUALAN MOTOR PADA PT. CAPELLA DINAMIK NUSANTARA CABANG MUKA KUNING,” 2018.
- [7] S. Anggraini, S. Defit, G. W. Nurcahyo, J. R. Lubuk, B. Padang, and S. Barat, “Analisis Data Mining Penjualan Ban Menggunakan Algoritma C4.5,” 2018.

- [8] N. Nafiiyah, "Algoritma." [Online]. Available: <https://warstek.com/algoritma/>
- [9] M. Kaufmann, "Decision Trees." [Online]. Available: [https://scikit-learn.org/stable/modules/tree.html#:~:text=Decision Trees \(DTs\) are a,as a piecewise constant approximation.](https://scikit-learn.org/stable/modules/tree.html#:~:text=Decision%20Trees%20(DTs)%20are%20a%20piecewise%20constant%20approximation.)
- [10] A. Wijaya, A. Mutatkin Bakti, T. Informatika, U. Bina Darma, J. A. Jendral Yani No, and P. Sumatera Selatan, "Implementasi Data Mining Menggunakan Algoritma C4," 2023.
- [11] E. Elisa, "PENERAPAN ALGORITMA C4.5 UNTUK MEMPREDIKSI PENJUALAN BARANG PADA PT BATAM BANGUN PRATHAMA," *JURNAL COMASIE*, 2022.
- [12] N. I. Purnama and L. P. Putri, "Analisis Penggunaan E-Commerce Di Masa Pandemi," 2021.
- [13] N. M. Maitri Priandewi, "PENGARUH KUALITAS PRODUK, DESAIN PRODUK, DAN PROMOSI TERHADAP KEPUTUSAN PEMBELIAN PRODUK DI E-COMMERCE LAZADA," *Jurnal Ilmiah Ilmu Manajemen dan Kewirausahaan Universitas Bina Bangsa*, vol. 1, 2021.