

Implementasi Algoritma C4.5 Untuk Memprediksi Penjualan Paket Internet

Armadani Simanjorang¹, Akim M.H Pardede², Siswan Syahputra³

^{1,2,3}STMIK Kaputama, Jl. Veteran No.4A-9A Binjai, 20714, Indonesia

ARTICLE INFORMATION

Received: September 09, 2024

Revised: September 12, 2024

Available online: Oktober 30, 2024

KEYWORDS

Algoritma C4.5,
Prediksi Penjualan,
Paket Internet,
Pohon Keputusan.

CORRESPONDENCE

Phone: +62 82361180280

E-mail: simanjorangarmadani@gmail.com

A B S T R A C T

Kebutuhan akan paket internet semakin meningkat seiring dengan perkembangan teknologi yang semakin pesat. Hal ini membuat perusahaan telekomunikasi seperti telkomsel menghadapi tantangan dalam memprediksi penjualan paket internet. Prediksi yang akurat dapat membantu perusahaan dalam menyusun strategi pemasaran dan penyediaan stok paket yang efektif. Penelitian ini dilakukan untuk memprediksi penjualan paket internet berdasarkan data penjualan dari PT. Golden Communication di kota Binjai. Proses penelitian ini menggunakan algoritma C4.5 untuk membentuk pohon keputusan (Decision Tree) yang dihasilkan perhitungan entropy dan gain dari berbagai variabel seperti Paket, Harga, Masa Aktif dan Terjual. Hasil dari penelitian ini menunjukkan bahwa algoritma C4.5 mampu memprediksi penjualan paket internet menggunakan aplikasi RapidMiner dengan tingkat akurasi sebesar 94,5 % dengan 200 data testing. Berdasarkan analisis yang dilakukan faktor-faktor yang paling berpengaruh terhadap penjualan paket internet adalah masa aktif, paket dan harga. Paket dengan masa aktif yang lebih pendek dan harga yang lebih murah cenderung lebih diminati oleh konsumen.

PENDAHULUAN

Dalam perkembangan teknologi saat ini, masyarakat telah memanfaatkan teknologi dalam kegiatan sehari-hari. Pemanfaatan teknologi yang telah diterapkan yaitu dalam bidang informasi, komunikasi, transportasi, pendidikan, kesehatan, konstruksi, pertanian, arsitektur dan *Artificial Intelligence (AI)*. Perkembangan teknologi khususnya pada komputer dan smartphone memerlukan akses jaringan internet dalam penggunaan situs *web*, *media social*, *e-mail*, *e-commerce* dan berbagai *platform* lainnya.

Untuk mengakses jaringan internet memerlukan kuota internet, jika sebuah perangkat tidak ada kuota internet maka perangkat tidak akan terhubung dengan internet. Oleh karena itu, berbagai *Internet Service Provider (ISP)* menyediakan kuota internet mulai dari kuota terkecil hingga kuota terbesar dengan masa aktif yang berbeda-beda. Saat ini, menjual kuota internet menjadi salah satu bisnis yang menjanjikan. Karena kuota internet sekarang menjadi kebutuhan yang sangat penting bagi masyarakat dalam memenuhi kebutuhan sehari-hari.

Dengan demikian, PT. Golden Communication adalah salah satu perusahaan saluran distribusi penjualan paket internet dan kartu perdana. Dalam penjualan paket internet yang mengalami kendala dalam penyediaan barang untuk periode waktu yang akan datang. Dimana setiap harinya permintaan paket internet sering terjadi perubahan yang tidak menentu. Penumpukan persediaan kuota internet dalam bentuk voucher maupun kartu perdana dengan masa aktif yang semakin hari semakin habis tidak sesuai dengan minat pembeli.

Oleh karena itu, penulis membuat suatu solusi dalam memprediksi penjualan paket internet menggunakan salah satu algoritma dalam data mining yaitu algoritma C4.5. Algoritma C4.5 adalah salah satu algoritma pohon keputusan yang populer dan efektif untuk masalah klasifikasi dan prediksi. Algoritma ini mampu mengatasi data yang kompleks dan dapat memberikan hasil yang dapat ditafsirkan.

METODE PENELITIAN

Metodologi penelitian adalah suatu cara atau proses yang digunakan dalam menyelesaikan sebuah penelitian. Metodologi ini mencakup berbagai teknik dan prosedur yang digunakan dalam proses pengumpulan, analisis, interpretasi dan penyajian data penelitian. Metodologi dalam penelitian ini digunakan untuk memprediksi penjualan paket internet dimasa yang akan datang berdasarkan data penjualan sebelumnya. Ada beberapa tahapan yang digunakan dalam memprediksi penjualan paket internet dalam program data mining yaitu sebagai berikut:

1. Identifikasi Masalah

Pada tahap ini, dilakukan untuk mengidentifikasi terhadap masalah yang ada dalam memprediksi penjualan paket internet. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang mempengaruhi penjualan paket internet dan menentukan metode yang tepat untuk memprediksi penjualan berdasarkan data historis penjualan.

2. Kajian Teori

Pada tahap ini, dilakukan tinjauan literature yang meliputi konsep dasar dari data mining, algoritma C4.5, teori tentang penjualan paket internet serta metode-metode prediksi yang telah digunakan pada penelitian sebelumnya. Tinjauan ini membantu dalam memahami konsep dan teori yang relevan dengan penelitian serta mendukung dasar pemilihan algoritma C4.5 sebagai metode yang akan diimplementasikan.

3. Pengumpulan Data

Tahap pengumpulan data melibatkan pengumpulan data historis penjualan paket internet dari salah satu penyedia layanan internet. Data yang dikumpulkan meliputi informasi penjualan harian, mingguan atau bulanan serta informasi lainnya yang dapat mempengaruhi prediksi penjualan paket internet.

4. Analisis data dilakukan untuk memahami pola dan tren dalam data penjualan yang telah dikumpulkan. Pada tahap ini, data dibersihkan dari anomali dan ketidakkonsistenan, kemudian dilakukan analisis untuk mengevaluasi hubungan antara berbagai variabel yang dapat mempengaruhi penjualan paket internet. Hasil analisis digunakan untuk menentukan atribut yang akan digunakan dalam model prediksi.

5. Implementasi Metode

Tahap implementasi metode melibatkan penerapan algoritma C4.5 untuk membuat model prediksi penjualan paket internet. Algoritma ini digunakan untuk membangun pohon keputusan yang dapat memprediksi penjualan berdasarkan atribut-atribut yang telah diidentifikasi pada tahap analisis data. Model yang dihasilkan diuji dan dioptimalkan untuk meningkatkan akurasi prediksi.

6. Pengujian Sistem

Tahap pengujian sistem dilakukan untuk mengevaluasi kinerja model prediksi yang telah dibangun. Pengujian dilakukan dengan membandingkan hasil prediksi dengan data *actual* penjualan dan menghitung metric-metric evaluasi seperti akurasi, presisi dan recall. Selain itu, dilakukan juga pengujian terhadap berbagai scenario untuk memastikan model dapat beradaptasi dengan berbagai kondisi pasar dan variasi data.

HASIL DAN PEMBAHASAN

1. Analisa Algoritma C4.5

Dalam analisa pengujian metode algoritma C4.5 dalam sistem prediksi yang digunakan diperlukan data sebagai *input* proses dan analisa. Berikut ini adalah tabel data pendukung penelitian.

Tabel 1. Data Pendukung Keputusan

No	Paket(GB)	Harga	Masa Aktif (Hari)	Terjual
1	2	Rp 10,400	3	502
2	3	Rp 12,800	5	1519
3	5	Rp 21,500	5	550
4	7	Rp 26,000	7	31
5	3	Rp 11,400	3	570
6	SA 4	Rp 20,000	30	4
7	BYU 7	Rp 15,000	30	8

No	Paket(GB)	Harga	Masa Aktif (Hari)	Terjual
8	3	Rp 11,400	3	550
9	3	Rp 12,800	5	2370
10	5	Rp 21,500	5	525
11	7	Rp 26,000	7	18
12	SA 30	Rp 75,000	30	5
13	BYU 14	Rp 38,000	30	14
14	BYU 20	Rp 48,000	30	3
15	BYU 20	Rp 48,000	30	8
16	3	Rp 11,400	3	550
17	3	Rp 12,800	5	1416
18	5	Rp 21,500	5	515
19	7	Rp 26,000	7	507
20	SA 35	Rp 62,000	30	2
21	3	Rp 11,400	3	600
22	2	Rp 10,400	3	44
23	3	Rp 12,800	5	1327
24	5	Rp 21,500	5	509
25	7	Rp 26,000	7	92
26	SA 4	Rp 20,000	30	4
27	BYU 7	Rp 15,000	30	5
28	3	Rp 11,400	3	16
29	2	Rp 10,400	3	600
30	3	Rp 12,800	5	1428
31	5	Rp 21,500	5	512
32	7	Rp 26,000	7	536
33	SA 35	Rp 62,000	30	1
34	BYU 7	Rp 15,000	30	9
35	3	Rp 11,400	3	72
36	3	Rp 12,800	5	1500
37	5	Rp 21,500	5	540
38	7	Rp 26,000	7	505
39	BYU 14	Rp 38,000	30	1
40	BYU 7	Rp 15,000	30	16
41	BYU 7	Rp 20,000	30	47
42	BYU 7	Rp 15,000	30	87
43	SA 4	Rp 20,000	30	5
44	SA 35	Rp 62,000	30	2
45	BYU 20	Rp 48,000	30	4
46	BYU 2	Rp 13,000	30	9
47	SA 50	Rp 75,000	30	60
48	SA 50	Rp 68,000	30	33

Selanjutnya melakukan transformasi data, data dilihat pada tabel 2.

Tabel 2. Transformasi Data

No.	Variabel	Transformasi	Nilai
-----	----------	--------------	-------

			Transformasi
1.	Harga(Rp)	$\leq 25,000$	Murah
		$>26,000-50,000$	Sedang
		$>50,000-75,000$	Mahal
2.	Masa Aktif	3 Hari	1
		5 Hari	2
		7 Hari	3
		30 Hari	4
3.	Terjual	<500	Rendah
		$>500-1000$	Sedang
		>1000	Tinggi

Setelah dilakukan pengelompokan untuk atribut paket, masa aktif, harga dan terjual maka dibuatlah tabel yang isi fieldnya menggunakan data yang sudah dikelompokkan. Data tersebut dapat dilihat pada tabel 3.

Tabel 3. Data Setelah Transformasi

No	Paket(GB)	Harga	Masa Aktif	Terjual
1	2	Murah	1	Sedang
2	3	Murah	2	Tinggi
3	5	Murah	2	Sedang
4	7	Sedang	3	Rendah
5	3	Murah	1	Sedang
6	SA 4	Murah	4	Rendah
7	BYU 7	Murah	4	Rendah
8	3	Murah	1	Sedang
9	3	Murah	2	Tinggi
10	5	Murah	2	Sedang
11	7	Sedang	3	Rendah
12	SA 30	Mahal	4	Rendah
13	BYU 14	Sedang	4	Rendah
14	BYU 20	Sedang	4	Rendah
15	BYU 20	Sedang	4	Rendah
16	3	Murah	1	Sedang
17	3	Murah	2	Tinggi
18	5	Murah	2	Sedang
19	7	Sedang	3	Sedang
20	SA 35	Mahal	4	Rendah
21	3	Murah	1	Sedang
22	2	Murah	1	Rendah
23	3	Murah	2	Tinggi
24	5	Murah	2	Sedang
25	7	Sedang	3	Rendah
26	SA 4	Murah	4	Rendah
27	BYU 7	Murah	4	Rendah
28	3	Murah	1	Rendah
29	2	Murah	1	Sedang
30	3	Murah	2	Tinggi
31	5	Murah	2	Sedang

No	Paket(GB)	Harga	Masa Aktif	Terjual
32	7	Sedang	3	Sedang
33	SA 35	Mahal	4	Rendah
34	BYU 7	Murah	4	Rendah
35	3	Murah	1	Rendah
36	3	Murah	2	Tinggi
37	5	Murah	2	Sedang
38	7	Sedang	3	Sedang
39	BYU 14	Sedang	4	Rendah
40	BYU 7	Murah	4	Rendah
41	BYU 7	Murah	4	Rendah
42	BYU 7	Murah	4	Rendah
43	SA 4	Murah	4	Rendah
44	SA 35	Mahal	4	Rendah
45	BYU 20	Sedang	4	Rendah
46	BYU 2	Murah	4	Rendah
47	SA 50	Mahal	4	Rendah
48	SA 50	Mahal	4	Rendah

a. Perhitungan Secara Manual

Untuk penyelesaian kasus pada algoritma C4.5 ada beberapa elemen yang harus dicari nilainya yaitu:

1. *Entropy(S)* adalah parameter yang digunakan untuk mengukur keberagaman setiap nilai attribute kriteria terhadap *decision attribute* (attribute keputusan) dalam sebuah kumpulan data. Rumus *Entropy* sebagai berikut :

$$Entropy(S) = -\sum_{i=1}^n p_i * \log_2(p_i)$$

Keterangan :

S = Jumlah sampel data (*Sampling*)

n = jumlah partisi S

pi = proporsi dari Sj terhadap S.

2. Gain (S,A) merupakan selisih nilai *entropy* total dikurangi nilai *entropy* masing-masing nilai setiap attribute kriteria dilaki nilai proporsi nilai *attribute* dibagi jumlah sampel data. Fungsi dari nilai gain adalah untuk mengukur efektivitas masing-masing attribute kriteria dalam mengklasifikasikan data. Pada algoritma C4.5 nilai Gain digunakan sebagai dasar pembentukan *node* atau akar dan cabang pohon keputusan. Rumus *Gain* adalah sebagai berikut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan:

S = Jumlah Kasus (*Sampling*)

A = Attribute

n = Jumlah Partisi S

|Si| = Jumlah kasus pada partisi ke-i

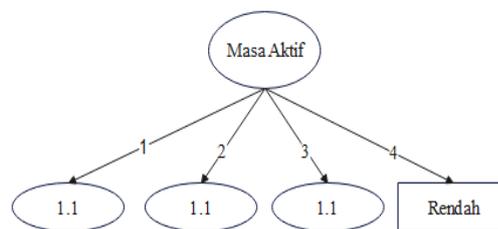
|S| = Jumlah kasus dalam S

Menghitung jumlah kasus, jumlah kasus untuk keputusan **Rendah, Sedang, Tinggi** dan *entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut **Paket, Masa Aktif dan Harga**. Selain itu melakukan perhitungan *Gain* untuk setiap atribut. Hasil perhitungan dapat dilihat pada tabel

Tabel 4. Perhitungan Node 1

Node	Variabel	Value	Jumlah Kasus(S)	Rendah (S1)	Sedang (S2)	Tinggi (S3)	Entropy	Gain
------	----------	-------	-----------------	-------------	-------------	-------------	---------	------

Node	Variabel	Value	Jumlah Kasus(S)	Rendah (S1)	Sedang (S2)	Tinggi (S3)	Entropy	Gain
1	Total		48	27	15	6	1.36631466	
	Banyak Paket							1.001527678
		2 GB	3	1	2	0	0	
		3 GB	12	2	4	6	1.459147917	
		5 GB	6	0	6	0	0	
		7 GB	6	3	3	0	0	
		SA 30 GB	1	1	0	0	0	
		SA 35 GB	3	3	0	0	0	
		SA 4 GB	3	3	0	0	0	
		BYU 7 GB	6	6	0	0	0	
		BYU 20 GB	3	3	0	0	0	
		BYU 14GB	2	2	0	0	0	
		SA 50 GB	2	2	0	0	0	
		BYU 2 GB	1	1	0	0	0	
	Harga							0.388292901
		MURAH	31	13	12	6	1.514356267	
		SEDAN G	11	8	3	0	0	
		MAHAL	6	6	0	0	0	
	Masa Aktif							1.366314657
		1	9	3	6	0	0	
		2	12	0	6	6	0	
		3	6	3	3	0	0	
		4	21	21	0	0	0	



Gambar 1. pohon Keputusan Node 1

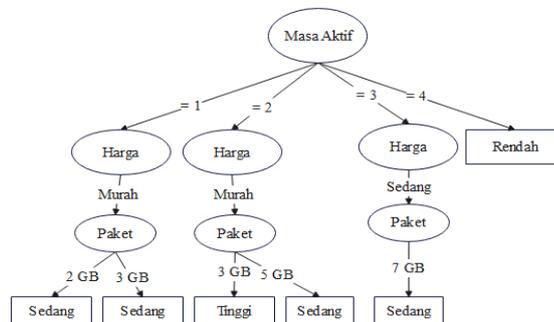
Selanjutnya, dilakukan perhitungan jumlah kasus **Rendah**, **Sedang**, **Tinggi** dan *entropy* dari semua kasus berdasarkan atribut atribut **Paket** dan **Harga**. Setelah itu lakukan perhitungan *Gain* untuk setiap atribut, hasil perhitungan dapat dilihat pada tabel 5.

Tabel 5. Perhitungan Node 1,1

N	Variabel	Value	Jumlah Kasus(S)	Rendah (S1)	Sedang (S2)	Tinggi (S3)	Entropy	Gain
1.1	Masa Aktif	1,2,3	27	6	15	6	1.4355205	
	Harga							0.363133343
		MURAH	21	3	12	6	1.37878349	
		SEDANG	6	3	3	0	0	
	Paket							0.787010319

N	Variabel	Value	Jumlah Kasus(S)	Rendah (S1)	Sedang (S2)	Tinggi (S3)	Entropy	Gain
		2 GB	3	1	2	0	0	
		3 GB	12	2	4	6	1.45914792	
		7 GB	6	3	3	0	0	

Dari hasil tabel 5 dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah **Paket** yaitu sebesar 0,787010319 dengan demikian, Paket dapat menjadi *node* akar. Dari hasil tersebut digambarkan pohon keputusan seperti pada gambar 3.



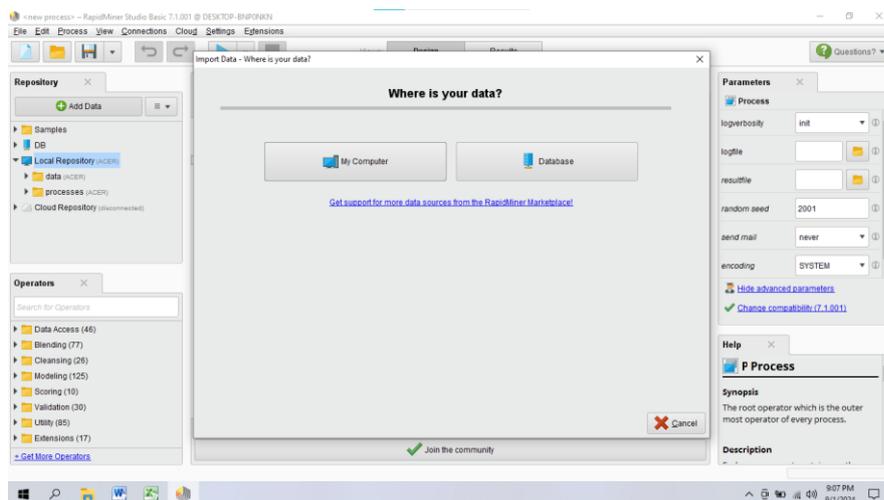
Gambar 2. Pohon Keputusan *Node* 1.1

3. Perhitungan Menggunakan *RapidMiner*

Pengujian akan dilakukan dengan memproses data input, proses prediksi menggunakan aplikasi *rapidminer*, yaitu sebagai berikut:

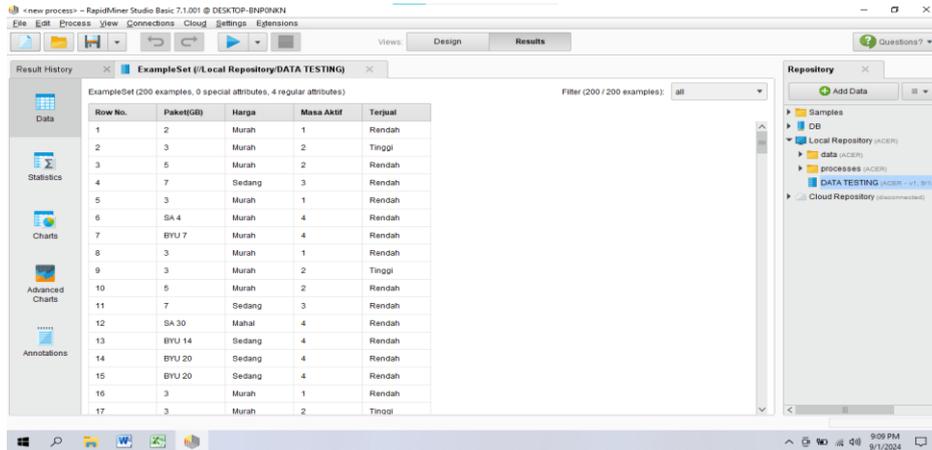
1. Proses Input data

Pilih file data yang akan diproses oleh aplikasi *RapidMiner*. Data yang akan di proses adalah data dengan nama file Paket Telkomsel.xlsx, berikut tampilan yang ada pada proses pengambilan data pada *RapidMiner*:



Gambar 3. Pengambilan Data Pada *RapidMiner*

Kemudian data penjualan paket akan muncul pada tampilan *RapidMiner*.

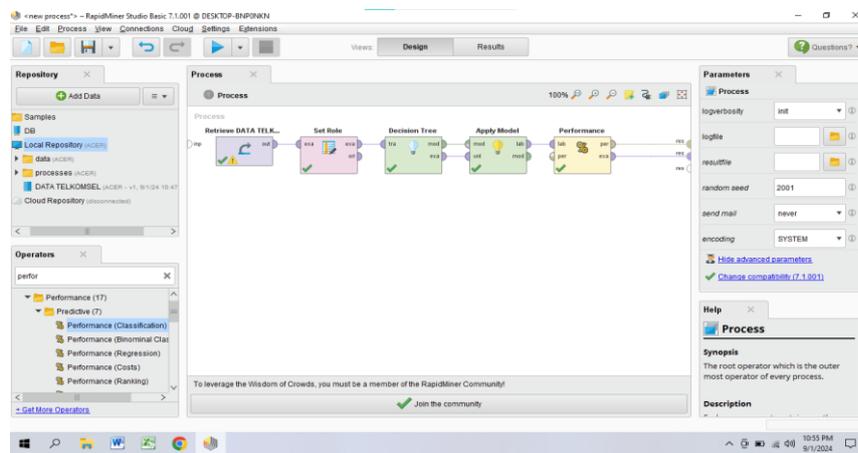


Row No.	Paket(GB)	Harga	Masa Aktif	Tertjual
1	2	Murah	1	Rendah
2	3	Murah	2	Tinggi
3	5	Murah	2	Rendah
4	7	Sedang	3	Rendah
5	3	Murah	1	Rendah
6	SA 4	Murah	4	Rendah
7	BYU 7	Murah	4	Rendah
8	3	Murah	1	Rendah
9	3	Murah	2	Tinggi
10	5	Murah	2	Rendah
11	7	Sedang	3	Rendah
12	SA 30	Mahal	4	Rendah
13	BYU 14	Sedang	4	Rendah
14	BYU 20	Sedang	4	Rendah
15	BYU 20	Sedang	4	Rendah
16	3	Murah	1	Rendah
17	3	Murah	2	Tinggi

Gambar 4. Tampilan Data Setelah Di Input

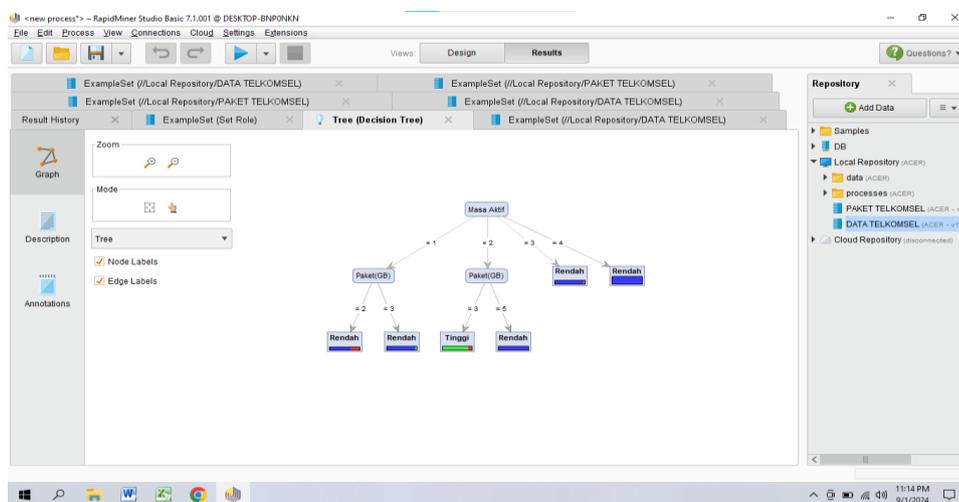
2. Rancangan Desain Proses Pada *RapidMiner*

Rancangan desain proses untuk pembentukan pohon keputusan pada *RapidMiner* dapat dilihat pada gambar dibawah.



Gambar 5. Desain Proses pada RapiMiner

Setelah dilakukan *running proses* untuk mendapatkan hasil berupa pohon keputusan. Berikut ini adalah pohon keputusan yang didapatkan.



Gambar 6. Hasil Pohon Keputusan (Decision Tree)

Rule yang dihasilkan dari pohon keputusan tersebut adalah sebagai berikut:

Tree

```

Masa Aktif = 1
| Paket (GB) = 2: Rendah {Rendah=15, Tinggi=0, Sedang=6}
| Paket (GB) = 3: Rendah {Rendah=23, Tinggi=1, Sedang=0}
Masa Aktif = 2
| Paket (GB) = 3: Tinggi {Rendah=0, Tinggi=23, Sedang=3}
| Paket (GB) = 5: Rendah {Rendah=26, Tinggi=0, Sedang=0}
Masa Aktif = 3: Rendah {Rendah=24, Tinggi=0, Sedang=1}
Masa Aktif = 4: Rendah {Rendah=78, Tinggi=0, Sedang=0}
    
```

Gambar 7. Rule Pohon Keputusan

- Jika Masa Aktif = 1(3 hari) dengan Paket = 2 GB maka Terjual = Rendah
 - Jika Masa Aktif = 1(3 hari) dengan Paket = 3 GB maka Terjual = Rendah
 - Jika Masa Aktif = 2 (5 hari) dengan Paket = 3 GB maka Terjual = Tinggi
 - Jika Masa Aktif = 2 (5 hari) dengan Paket = 5 GB maka Terjual = Rendah
 - Jika Masa Aktif = 3 (7 hari) maka Terjual = Rendah
 - Jika Masa Aktif = 4 (30 hari) maka Terjual = Rendah
4. Evaluasi Model

Setelah proses prediksi penjualan paket internet menggunakan algoritma C4.5 pada *RapidMiner* selanjutnya yang akan dilakukan adalah evaluasi model dengan menghitung *accuracy*, *precision*, *recall* dan *error rate* pada data dengan *confusion matrix*.

Tabel 6. Confusion Matrix

Actual	Predicted		
	Tinggi	Sedang	Rendah
Tinggi	23	3	0
Sedang	0	0	0
Rendah	1	7	166

Dari tabel *Confusion Matrix* diperoleh:

- True Positif* = 166 + 23 = 189
- True Negative* = 0
- False Negative* = 0
- False Positive* = 1+3 +7 = 11

Maka diperoleh perhitungan *accuracy*, *precision*, *recall* dan *error rate*, yaitu sebagai berikut :

- $Accuracy = \frac{TN+TP}{TN+TP+FN+FP} = \frac{0+189}{0+189+0+11} = 0,945 \times 100 \% = 94,5 \%$
- $Precision = \frac{TP}{FP+TP} = \frac{189}{0+189} = 0,945 \times 100\% = 94,5 \%$
- $Recall = \frac{TP}{TP+FN} = \frac{189}{189+0} = 1 \times 100\% = 100 \%$
- $Erorr Rate = \frac{FN+FP}{TN+TP+FN+FP} = \frac{11+0}{0+189+0+11} = 0,55 \times 100 \% = 5,5 \%$

Hasil *Performance Testing* pada *RapidMiner Performance* meliputi hasil perhitungan *accuracy*. Hasil perhitungan pada *rapidminer* dapat dilihat pada gambar 6.

Table View Plot View

accuracy: 94.50%

	true Rendah	true Tinggi	true Sedang	class precision
pred. Rendah	166	1	7	95.40%
pred. Tinggi	0	23	3	88.46%
pred. Sedang	0	0	0	0.00%
class recall	100.00%	95.83%	0.00%	

Gambar 8. Evaluasi Model Pada RapidMiner

KESIMPULAN

Setelah dilakukan pengujian data penjualan paket internet menggunakan *RapidMiner* pada PT. Golden Communication sebanyak 200 data, maka diperoleh pohon keputusan dengan *rule* terbaik untuk penjualan paket internet dengan terjual Tinggi yaitu Paket 3 GB dengan Masa Aktif = 2 (5 hari). Berdasarkan hasil evaluasi model pada penelitian prediksi penjualan paket internet menggunakan Algoritma C4.5 diperoleh nilai *accuracy* sebesar 94,5 %. Implementasi algoritma C4.5 terbukti efektif dalam memprediksi penjualan paket internet di PT. Golden Communication. Algoritma ini mampu membentuk pohon keputusan yang akurat berdasarkan atribut-atribut seperti Paket, Masa Aktif, Harga, Dan Terjual. Pohon keputusan yang dihasilkan memberikan insight yang dapat membantu pengelola dalam menentukan strategi penjualan di masa mendatang. Model prediksi yang dibangun menggunakan algoritma C4.5 menunjukkan tingkat akurasi yang cukup tinggi dalam memprediksi penjualan paket internet. Hal ini menandakan bahwa model ini dapat diandalkan untuk membantu perusahaan dalam mengelola persediaan dan mengoptimalkan penjualan berdasarkan prediksi yang dihasilkan.

REFERENSI

- Buulolo, E. (2020). *Data Mining : Untuk Perguruan Tinggi*.
- Helmud, E. (2016). Pemilihan Paket Internet Android Pada Operator Telepon Gsm Menggunakan Metode Analytical Hierarchy Process (Ahp). *Jurnal Sistem Informasi (JSI)*, 8(1), 1. <http://ejournal.unsri.ac.id/index.php/jsi/index>
- Kafil, M. (2019). Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 3(2), 59–66. <https://doi.org/10.36040/jati.v3i2.860>
- Kusrini, & Lutfi, E. T. (2009). *Algoritma data mining*.
- Leonardi, M., Emilda, R., Katrin, I., & Yulianto, A. (2021). Prediksi Penjualan Produk Rokok Pada PT. Indomarco Prismatama Menggunakan Algoritma C4.5. *Paradigma - Jurnal Komputer Dan Informatika*, 23(2). <https://doi.org/10.31294/p.v23i2.11151>
- Marlina, D., & Bakri, M. (2021). Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4.5. *Jurnal Teknologi Dan Sistem Informasi (JTISI)*, 2(1), 23–28.
- Nas, C. (2021). Data Mining Prediksi Minat Calon Mahasiswa Memilih Perguruan Tinggi Menggunakan Algoritma C4.5. *Jurnal Manajemen Informatika (JAMIKA)*, 11(2), 131–145. <https://doi.org/10.34010/jamika.v11i2.5506>
- Rosaly, R., & Prasetyo, A. (2020). Flowchart Beserta Fungsi dan Simbol-Simbol. *Journal of Chemical Information and Modeling*, 2(3), 5–7.
- Sari, R. M. (2015). Prediksi Data Anggaran Pendapatan Belanja Daerah Menggunakan Algoritma K-Means. *Sains Dan Teknologi Informasi*, 1(2), 1–6.
- Wahyudi, M. D. (2023). Penerapan Data Mining Dengan Algoritma C4. 5 Dalam Prediksi Penjualan Buku. *Jurnal Teknorama (Informatika Dan..., 1(1), 1–6*. <https://jurnal.stikomelrahma.ac.id/index.php/teknorama/article/view/1%0Ahttps://jurnal.stikomelrahma.ac.id/index.php/teknorama/article/download/1/1>