

Teknik Klasifikasi C4.5 Dalam Menentukan Faktor Utama Kepuasan Nasabah Terhadap Pelayanan Klaim di PT Asuransi Central Asia Pematangsiantar

Muhammad Syahputra, Agus Perdana Windarto, Riki Winanjaya*

Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: muhammedsyahputra@gmail.com

Abstrak—Klaim asuransi dapat dilakukan sesuai dengan ketentuan penting dalam pengajuan klaim, seperti klaim sesuai dengan yang tertera dalam polis, polis masih berlaku (*inforce*), polis tidak dalam masa tunggu dan klaim termasuk dalam pertanggungan. *Data Mining* merupakan proses yang dibantu oleh komputer untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi dan pengetahuan. *Decision Tree* adalah salah satu dari bagian *Data Mining*, dengan algoritma C4.5 merupakan salah satu metode klasifikasi yang menggunakan representasi pohon keputusan dimana setiap node mempresentasikan atribut, cabangnya mempresentasikan nilai dari atribut, daun mempresentasikan kelas. dan dapat digunakan untuk mencari faktor dominan untuk mencari suatu keputusan, salah satunya menentukan faktor utama kepuasan nasabah terhadap pelayanan klaim di PT. Asuransi Central Asia Pematangsiantar. Yang diharapkan hasilnya nanti dapat memberikan kontribusi yang besar kepada PT. Asuransi Central Asia Pematangsiantar dalam memberikan pelayanan terbaik kepada nasabah yang ada.

Kata Kunci: Data Mining; Algoritma C4.5; Klaim Asuransi; Kepuasan Nasabah

Abstract—Insurance claims can be made in accordance with important provisions in claim submission, such as claims in accordance with those stated in the policy, the policy is still valid (*inforce*), the policy is not in the waiting period and the claim is included in the coverage. Data Mining is a computer-aided process to explore and analyze large amounts of data sets and extract information and knowledge. Decision Tree is one of the Data Mining sections, with C4.5 algorithm is one of the classification methods that use decision tree representations where each nodes present attributes, branches represent the values of attributes, leaves represent classes. and can be used to find dominant factors to find a decision, one of which determines the main factors of customer satisfaction with service claims at PT. Insurance Central Asia Pematangsiantar. It is hoped that the results will later be able to contribute greatly to PT. Insurance Central Asia Pematangsiantar in providing the best service to existing customers.

Keywords: Data Mining; C4.5 Algorithm; Insurance Claims; Customer Satisfaction

1. PENDAHULUAN

Data Mining secara umum proses yang menggunakan berbagai teknik seperti teknik statistik, matematika, kecerdasan buatan untuk mengekstrak dan mengidentifikasi berbagai informasi yang bermanfaat untuk menghasilkan pengetahuan yang terkait dengan data yang digunakan dalam skala besar. *Data Mining* termasuk dalam proses *Knowledge Discovery Database* yang merupakan proses yang dibantu oleh komputer untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi dan pengetahuan. *Data Mining* terdiri dari beberapa fase yaitu Estimasi, Klustering, Prediksi, Klasifikasi dan Asosiasi. Salah satu Data Mining yang populer digunakan adalah fase Klasifikasi. *Data Mining* fase Klasifikasi memiliki salah satu metode yang populer menurut Xindong Wu dan Vipin Kumar dalam bukunya yang berjudul *The Top Ten Algorithms in Data Mining*, yaitu *Decision Tree*. *Decision Tree* dengan algoritma C4.5 merupakan salah satu metode klasifikasi yang menggunakan representasi pohon keputusan dimana setiap node mempresentasikan atribut, cabangnya mempresentasikan nilai dari atribut, dan daun mempresentasikan kelas (Sembiring, Sibuea & Septa, 2018). Algoritma C4.5 diperkenalkan oleh J. Rose Quinlan (1996) sebagai versi perbaikan dari *ID3 (Interactive Dichotomiser 3)* dan dapat digunakan untuk mencari faktor dominan untuk mencari suatu keputusan, salah satunya menentukan faktor utama kepuasan nasabah terhadap pelayanan klaim di PT. Asuransi Central Asia Pematangsiantar.

Pelayanan klaim asuransi merupakan permintaan resmi yang dilakukan oleh pihak tertanggung (nasabah) sehubungan dengan kontrak perjanjian antara pihak asuransi dan pihak tertanggung yang mengikatkan diri untuk menjamin pembayaran ganti rugi oleh penanggung (asuransi) ketika pihak tertanggung mengalami kejadian musibah yang diderita. Klaim asuransi dapat dilakukan sesuai dengan ketentuan penting dalam pengajuan klaim, seperti klaim sesuai dengan yang tertera dalam polis, polis masih berlaku (*inforce*), polis tidak dalam masa tunggu dan klaim termasuk dalam pertanggungan. Terdapat tiga tahapan dalam klaim asuransi, yaitu : *Notification*, *Investigation* dan *Submission*. Pihak tertanggung mengirimkan dokumen pendukung klaim yang diminta oleh penanggung. Penanggung melakukan pemeriksaan kesesuaian dokumen kepada polis, kelengkapan dokumen yang diminta oleh penanggung dan mengirimkan kepada pihak penanggung. Terdapat faktor yang mempengaruhi kepuasan nasabah dalam pelayanan klaim asuransi. Pelayanan yang diberikan oleh pihak penanggung diharuskan pelayanan yang terbaik untuk kepuasan pihak tertanggung dalam pengajuan pelayanan klaim. Dari berbagai faktor yang diberikan pihak penanggung terdapat faktor utama kepuasan pelayanan klaim yang dilakukan pihak tertanggung. Dengan adanya berbagai faktor yang digunakan, pihak penanggung tidak mengetahui faktor utama dalam kepuasan pelayanan nasabah di PT. Asuransi Central Asia Pematangsiantar. Faktor utama tersebut yang dijadikan pedoman untuk meningkatkan pelayanan klaim asuransi dengan tidak menjadikan faktor lainnya sebagai faktor utama dari Asuransi PT. Central Asia Pematangsiantar.

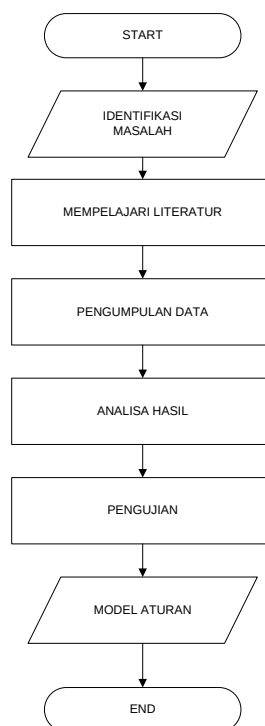
Permasalahan yang tertera diatas, penulis melakukan teknik klasifikasi dalam menentukan faktor utama kepuasan nasabah terhadap pelayanan klaim di PT. Asuransi Central Asia Pematangsiantar menggunakan *Data Mining* Algoritma C4.5. Data yang digunakan dengan memberikan kuesioner kepada nasabah yang telah melakukan pengajuan klaim asuransi di PT. Asuransi Central Asia Pematangsiantar.

Menganalisa faktor utama kepuasan nasabah dalam pelayanan klaim, penelitian ini dilandasi oleh penelitian sebelumnya oleh (Rohman, Suhartono & Supritanto, 2017), pada studi kasus : Penerapan Algoritma C4.5 Bebas *Adaboost* untuk Prediksi Penyakit Jantung dengan kesimpulan Algoritma C4.5 berbasis *Adaboost* memberikan pemecahan untuk permasalahan penyakit jantung lebih akurat dengan nilai *accuracy* 92,24% dengan nilai AUC adalah 0.982 dan penelitian sebelumnya oleh (Haryati, Sudarsono & Suryana, 2015), pada studi kasus : Implementasi Data Mining untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus : Universitas Dehasen Bengkulu) dengan kesimpulan Algoritma C4.5 lebih akurat dibandingkan analisa yang dilakukan oleh analis mahasiswa dengan hal ini dibuktikan dengan hasil evaluasi penelitian bahwa Algoritma C4.5 mampu menganalisa tingkat ketepatan waktu mahasiswa menyelesaikan masa studinya.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Rancangan atau model penelitian disajikan dalam rancangan *flowchart* pada gambar 1 sebagai berikut:



Gambar 1. Rancangan Penelitian

Pada gambar 1. menjelaskan rancangan penelitian yang dilakukan untuk menentukan Faktor Utama Kepuasan Nasabah Terhadap Pelayanan Klaim di PT. Asuransi Central Asia Pematangsiantar yang terdiri dari:

- Identifikasi Masalah**
Mengidentifikasi masalah yang terkait dalam menentukan faktor utama kepuasan nasabah berdasarkan parameter yang digunakan dalam penelitian.
- Mempelajari Literatur**
Penelitian ini dilandasi rujukan atau referensi yang terkait guna mendapatkan informasi pendukung dalam penelitian.
- Pengumpulan data**
Pengumpulan data dilakukan dengan melakukan teknik wawancara dengan wakil kepala sekolah untuk menentukan alternatif dan teknik kuesioner kepada nasabah pada PT. Asuransi Central Asia Pematangsiantar waktu pengumpulan data dan pengolahan data dilakukan selama 2 Minggu dengan waktu dimulai tanggal 24 Februari s/d 9 Maret 2020
- Analisa Data**
Proses yang dilakukan untuk mencari faktor utama kepuasan nasabah terhadap pelayanan klaim.
- Pengujian**
Penulis melakukan pengujian dengan menggunakan aplikasi *Ms. Excel* dan *RapidMiner*.
- Model Aturan**

Hasil yang diberikan oleh sistem dan analisa penulis dapat memberikan aturan model faktor utama kepuasan nasabah terhadap pelayanan klaim di PT. Asuransi Central Asia Pematangsiantar.

2.2 Analisa Data

Dalam penelitian, data yang digunakan akan diolah dari hasil kuesioner yang diberikan kepada nasabah asuransi PT. Asuransi Central Asia Pematangsiantar dengan sampel sebanyak 50 data. Data yang digunakan sebagai parameter terdiri dari:

- Lama Menjadi Nasabah
- Tangible
- Reliability
- Assurance
- Responsive

Setiap faktor terdiri dari pertanyaan yang diajukan ke nasabah yang telah melakukan pelayanan klaim di PT. Asuransi Central Asia Pematangsiantar. Kuesioner yang telah dilakukan selanjutnya mencari rata-rata dari faktor yang digunakan. Data yang digunakan menggunakan jenis statistik *deskriptif* dengan kepada nasabah PT. Asuransi Central Asia Pematangsiantar. Kuesioner yang diberikan menggunakan *linker 5* yang terdiri dari SB (Sangat Baik), B (Baik), CB (Cukup Baik), KB (Kurang Baik), dan SKB (Sangat Kurang Baik). Data diolah menggunakan *RapidMiner* dengan operator *Performance* yang berfungsi sebagai validasi dan reabilitas data untuk mencari keakuratan data. Data yang akurat, dilakukan pengolahan data untuk mencari hasil dari masalah penelitian dengan menggunakan *RapidMiner* dan diambil keputusan dari hasil. Berikut sampel data yang digunakan sebanyak sampel 50 data pada tabel 1:

Tabel 1. Data Sampel

No	Alternatif	Lama Nasabah	Tangible	Reliability	Assurance	Responsive	Tanggapan
1	A1	L	SB	SB	SB	SB	Puas
2	A2	L	B	B	B	B	Puas
3	A3	BR	B	B	B	B	Puas
4	A4	CL	B	SB	B	SB	Puas
5	A5	L	B	B	B	B	Puas
6	A6	L	B	B	C	C	Tidak Puas
7	A7	L	SB	SB	B	C	Puas
8	A8	L	C	B	C	C	Tidak Puas
9	A9	BR	B	C	SB	B	Puas
10	A10	BR	B	C	SB	B	Puas
11	A11	BR	B	C	B	B	Puas
12	A12	CL	B	B	C	C	Puas
13	A13	CL	B	B	B	C	Puas
14	A14	CL	B	B	B	C	Puas
15	A15	CL	B	B	B	SB	Puas
16	A16	CL	B	B	B	SB	Puas
17	A17	L	B	B	C	SB	Puas
18	A18	L	B	B	B	C	Puas
19	A19	L	B	C	C	C	Tidak Puas
20	A20	L	B	B	C	C	Tidak Puas
21	A21	BR	C	B	B	C	Puas
22	A22	L	B	SB	SB	B	Puas
23	A23	L	B	SB	SB	B	Puas
24	A24	L	B	SB	B	B	Puas
25	A25	L	B	C	C	C	Tidak Puas
26	A26	BR	B	B	B	B	Puas
27	A27	CL	SB	SB	SB	C	Puas
28	A28	BR	SB	B	B	C	Puas
29	A29	L	SB	B	SB	SB	Puas
30	A30	BR	SB	B	SB	B	Puas
31	A31	BR	SB	SB	SB	B	Puas
32	A32	BR	SB	B	SB	SB	Puas
33	A33	CL	SB	SB	SB	B	Puas
34	A34	L	SB	SB	SB	SB	Puas
35	A35	BR	SB	B	B	C	Puas
36	A36	BR	SB	B	B	B	Puas
37	A37	CL	B	B	B	B	Puas

No	Alternatif	Lama Nasabah	Tangible	Reliability	Assurance	Responsive	Tanggapan
38	A38	BR	B	B	B	B	Puas
39	A39	CL	B	B	B	B	Puas
40	A40	L	SB	SB	SB	B	Puas
41	A41	CL	SB	SB	SB	SB	Puas
42	A42	CL	C	B	C	B	Puas
43	A43	BR	C	B	C	C	Tidak Puas
44	A44	CL	B	C	B	C	Tidak Puas
45	A45	BR	B	C	C	B	Tidak Puas
46	A46	BR	B	C	B	C	Puas
47	A47	BR	C	C	C	C	Tidak Puas
48	A48	BR	C	B	C	B	Tidak Puas
49	A49	CL	B	SB	B	SB	Puas
50	A50	L	SB	B	SB	B	Puas

2.3 Algoritma C4.5

Algoritma C4.5 adalah salah satu algoritma untuk mengubah fakta yang besar menjadi pohon keputusan (decision tree) yang merepresentasikan aturan (rule). Tujuan dari pembentukan pohon keputusan dalam algoritma C4.5 adalah untuk mempermudah dalam penyelesaian permasalahan (Faisal, 2019).

Kriteria yang paling banyak digunakan untuk memilih fitur sebagai pemecah dalam Algoritma C4.5 adalah *Rasio Gain*, yang diformulasikan oleh persamaan 1 berikut (Gerardo et al., 2019):

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

Keterangan :

S : Himpunan Kasus

A : Atribut

N : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

Sedangkan perhitungan nilai *Entropy* dapat dilihat pada persamaan 2 :

$$Entropy(A) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (2)$$

Keterangan :

S : Himpunan Kasus

A : Fitur

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

3. HASIL DAN PEMBAHASAN

3.1 Pengolahan Data

Permasalahan yang sudah dianalisa, selanjutnya penulis menentukan faktor penyebab penyebab Kepuasan Nasabah Terhadap Pelayanan Klaim di PT. Asuransi Central Asia dengan menggunakan *Data Mining* algoritma C4.5. Algoritma C4.5 dilakukan dengan merancang atau membentuk pohon keputusan (*decision tree*) dengan memilih atribut pada variabel data sebagai akar (*node*), membuat cabang untuk setiap nilai, membagi kasus dalam cabang, dan mengulangi proses sampai semua cabang memiliki kelas yang sama. Perhitungan *Entropy* dan *Gain* dilakukan untuk membentuk suatu keputusan. Membuat pohon keputusan terlebih dahulu menghitung jumlah kasus untuk keputusan Puas dan jumlah kasus keputusan Tidak Puas. Berikut ini langkah-langkah dalam membentuk pohon keputusan dengan menggunakan Algoritma C4.5 :

- Menghitung *Entropy* dari total data. Dari data penelitian diketahui jumlah kasus ada 50 *record*, diketahui status Puas ada 40 *record* dan status Tidak Puas ada 10 *record* sehingga *Entropy* nilai akar total yang didapat:

$$\begin{aligned} Entropy(S) &= \sum_{i=1}^n - p_i * \log_2 p_i \\ &= (-40/50 * \log_2 (40/50)) + (-10/50 * \log_2 (10/50)) \\ &= 0,7219 \end{aligned}$$

- Menghitung Menghitung nilai *Entropy* pada atribut Lama Nasabah dengan kategori Lama, jumlah kasus terdapat 18 *record*, terdapat Puas sebanyak 13 *record* dan status Tidak Puas sebanyak 5 *record*.

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

$$= (-13/18 \cdot \log_2 (13/18)) + (-5/18 \cdot \log_2 (5/18))$$

$$= 0,6840$$

- c. Seterusnya perhitungan *Entropy* pada atribut Cukup Lama sampai dengan kategori baru. Selanjutnya menghitung nilai *Gain* pada Atribut Lama Nasabah dengan kategori Lama terdapat 18 *record*, kategori Cukup Lama terdapat 14 *record*, dan kategori Baru terdapat 18 *record*

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

$$= 0,7219 - (((18/50)*0,8524)+((14/50)*0,3712)+((18/50)*0,7642))$$

$$= 0,0360$$

- d. Selanjutnya Menghitung *SplitInfo* pada atribut Lama Nasabah berdasarkan jumlah kasus pada Lama Nasabah.

$$SplitInfo = \sum_{i=1}^k p(v_i | s) \log_2 p(v_i | s)$$

$$= -((18/50.\log_2 (18/50)) + (14/50.\log_2 (14/50) + (18/50.\log_2 (18/50))))$$

$$= 1,5755$$

- e. Kemudian menghitung Nilai *GainRatio* Lama Nasabah sebagai nilai yang digunakan sebagai pemecahan *Node* pada pohon keputusan.

$$RasioGain(s,j) = \frac{Gain(s,j)}{SplitInfo(s,j)}$$

$$= 0,0360 / 1,5755$$

$$= 0,0229$$

Hasil dari perhitungan nilai *Entropy*, nilai *Gain*, *SplitInfo* dan *GainRatio* untuk tiap atribut dapat dilihat pada tabel 2 sebagai berikut:

Tabel 2. Hasil Perhitungan *Gain Node 1*

No	Atribut	Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain	SPLIT INFO	GAIN RATIO
1	Total	50	40	10	0,7219			
	Lama Nasabah					0,0360	1,5755	0,0229
	Lama	18	13	5	0,8524			
	Cukup Lama	14	13	1	0,3712			
2	Baru	18	14	4	0,7642			
	Tangible					0,1851	1,0712	0,1728
	Sangat Baik	15	15	0				
	Baik	29	23	6	0,7355			
3	Cukup	6	2	4	0,9183			
	Reability					0,1644	1,4190	0,1159
	Sangat Baik	13	13	0				
	Baik	28	23	5	0,6769			
4	Cukup	9	4	5	0,9911			
	Assurance					0,4085	1,5306	0,2669
	Sangat Baik	15	15	0				
	Baik	23	22	1	0,2580			
5	Cukup	12	3	9	0,8113			
	Responsive					0,1718	1,5161	0,1133
	Sangat Baik	10	10	0				
	Baik	22	20	2	0,4395			
	Cukup	18	10	8	0,9911			

Dari Tabel 2 diketahui nilai *GainRatio* terbesar adalah atribut *Assurance* dengan nilai 0,2669 kategori Baik dengan nilai *Entropy* 0,2580, kategori Cukup dengan nilai 0,8133 dan dijadikan simpul akar untuk mengulangi perhitungan nilai *Entropy* dan nilai *Gain* selanjutnya. Hasil Node 1.1 dapat dilihat pada tabel 3 berikut:

Tabel 3. Hasil Perhitungan *Gain Node 1.1*

No	Atribut	Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain	Split Info	Gain Ratio
1	Assurance - Baik	23	22	1	0,2580			
	Lama Nasabah					0,0611	1,5380	0,0397
	Lama	5	5	0				
	Cukup Lama	9	8	1	0,5033			
2	Baru	9	9	0				
	Tangible					0,0158	0,7310	0,0216

3	Reability	Sangat Baik	4	4	0	0,3095	0,1382	1,1864	0,1165
		Baik	18	17	1				
		Cukup	1	1	0				
5	Responsive	Sangat Baik	4	4	0	0,9183	0,0611	1,4910	0,0410
		Baik	16	16	0				
		Cukup	3	2	1				
		Sangat Baik	4	4	0	0,5033			
		Baik	10	10	0				
		Cukup	9	8	1				

Dari tabel 3 diketahui nilai *GainRatio* terbesar adalah atribut *Reability* dengan nilai 0,1165 kategori Cukup dengan nilai *Entropy* 0,9183, dan dijadikan simpul akar untuk mengulangi perhitungan nilai *Entropy* dan nilai *Gain* selanjutnya. dapat dilihat pada tabel 4 berikut:

Tabel 4. Hasil Perhitungan *Gain Node 1.2*

No	Atribut	Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain	Split Info	Gain Ratio
1	Assurance - Cukup	12	3	9	0,8113	0,4863	1,4591	0,3333
	Lama Nasabah							
	Lama	6	1	5	0,6500			
2	Tangible	Cukup Lama	2	2	0	0,0070	0,7776	0,0090
		Baru	4	0	4			
		Sangat Baik	0	0	0			
3	Reability	Baik	7	2	5	0,8631	0,1750	0,9183
		Cukup	5	1	4	0,7219		
		Sangat Baik	0	0	0	0,9544	0,2193	1,1887
5	Responsive	Baik	8	3	5			
		Cukup	4	0	4			
		Sangat Baik	1	1	0	0,9183	0,6500	1,0000
		Baik	3	1	2			
		Cukup	8	1	7			

Dari tabel 4 diketahui nilai *GainRatio* terbesar adalah atribut Lama Nasabah dengan nilai 0,3333 kategori Lama dengan nilai *Entropy* 0,6500, dan dijadikan simpul akar untuk mengulangi perhitungan nilai *Entropy* dan nilai *Gain* selanjutnya, dapat dilihat pada tabel 5 berikut :

Tabel 5. Hasil Perhitungan *Gain Node 1.2.1*

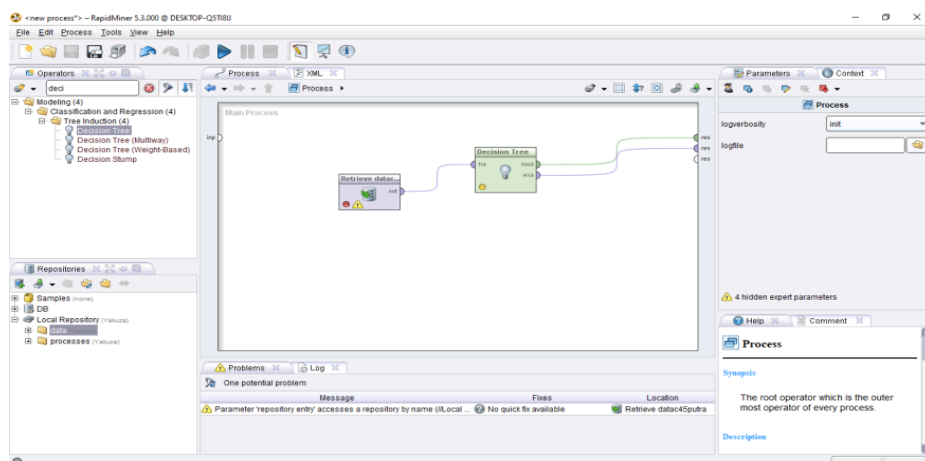
No	Atribut	Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain	Split Info	Gain Ratio
2	Assurance - Cukup, Lama Nasabah - Lama	6	1	5	0,6500	0,0484	0,2630	0,1841
	Tangible							
	Sangat Baik	0	0	0				
3	Reability	Baik	5	1	4	0,7219	0,1092	0,9183
		Cukup	1	0	1			
		Sangat Baik	0	0	0	0,8113	0,6500	1,0000
5	Responsive	Baik	4	1	3			
		Cukup	2	0	2			
		Sangat Baik	1	1	0	0,9183	0,6500	1,0000
		Baik	0	0	0			
		Cukup	5	0	5			

Dari tabel 5 diketahui nilai *GainRatio* terbesar adalah atribut *Responsive* dengan nilai 1,0000, Artinya bahwa *Node* 1.2.1 berhenti. Dengan demikian maka didapat bentuk pohon keputusan atau *Rule* model yang didapatkan adalah sebagai berikut :

- Jika assurance = baik, dan reliability = baik maka hasilnya Puas { Puas = 16 dan Tidak Puas = 0}
- Jika assurance = baik, reliability = cukup, dan Lama nasabah = baru maka hasilnya Puas {Puas = 2 dan Tidak Puas = 0}
- Jika assurance = baik, reliability = cukup, dan Lama nasabah = cukup lama maka hasilnya Tidak Puas {Puas = 0 dan Tidak Puas = 1}
- Jika assurance = baik dan reliability = sangat baik maka hasilnya Puas {Puas = 4 dan Tidak Puas = 0}
- Jika assurance = cukup dan Lama nasabah = baru maka hasilnya Tidak Puas {Puas = 0 dan Tidak Puas = 4}
- Jika assurance = cukup dan Lama nasabah = cukup lama maka hasilnya Puas {Puas = 2 dan Tidak Puas = 0}
- Jika assurance = cukup, Lama nasabah = lama, dan responsive = cukup maka hasilnya Tidak Puas {Puas = 0 dan Tidak Puas = 5}
- Jika assurance = cukup, Lama nasabah = lama, dan responsive = sangat baik maka hasilnya Puas {Puas = 1 dan Tidak Puas = 0}
- Jika assurance = sangat baik maka hasilnya Puas {Puas = 15 dan Tidak Puas = 0}

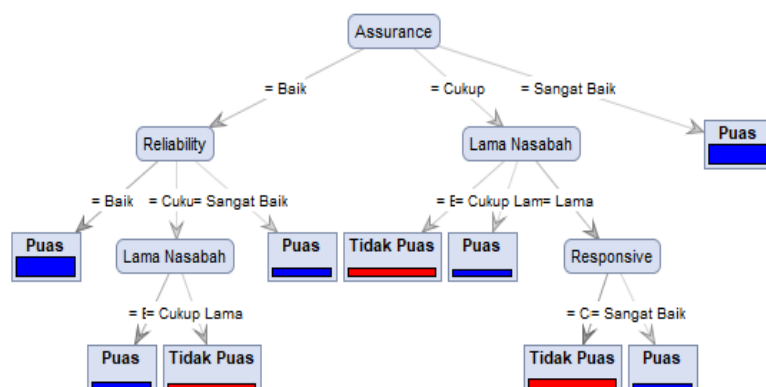
3.1 Implementasi Algoritma C4.5

Untuk membentuk model aturan penelitian, terlebih dahulu *importing* data dari data yang sudah diolah di *Microsoft Excel 2010* menggunakan perintah *Import Excel Sheet* pada *RapidMiner* agar bisa diakses menggunakan *software RapidMiner*. Atribut yang digunakan sebagai label adalah tingkat faktor Utama Kepuasan Nasabah Terhadap Pelayanan Klaim di PT. Asuransi Central Asia Pematangsiantar.



Gambar 2. Menghubungkan *Input* Dan *Output* Setiap Opertaor

Setelah dilakukan perhitungan dan pengujian pada data dengan menggunakan Algoritma C4.5 dan telah dilakukan penghubungan *port-port* dari *operator decision tree*, *operator apply model* dan *operator performance* maka akan menampilkan hasil :



Gambar 3. Tampilan *Decision Tree*

Gambar 3 diatas menghasilkan aturan atau *rule* yang dilakukan di *RapidMiner* menu *Decision Tree* pada tab *Text View* seperti gambar 4 berikut :

Tree

```

Assurance = Baik
| Reliability = Baik: Puas {Puas=16, Tidak Puas=0}
| Reliability = Cukup
| | Lama Nasabah = Baru: Puas {Puas=2, Tidak Puas=0}
| | Lama Nasabah = Cukup Lama: Tidak Puas {Puas=0, Tidak Puas=1}
| Reliability = Sangat Baik: Puas {Puas=4, Tidak Puas=0}
Assurance = Cukup
| Lama Nasabah = Baru: Tidak Puas {Puas=0, Tidak Puas=4}
| Lama Nasabah = Cukup Lama: Puas {Puas=2, Tidak Puas=0}
| Lama Nasabah = Lama
| | Responsive = Cukup: Tidak Puas {Puas=0, Tidak Puas=5}
| | Responsive = Sangat Baik: Puas {Puas=1, Tidak Puas=0}
Assurance = Sangat Baik: Puas {Puas=15, Tidak Puas=0}

```

Gambar 4. Tampilan Aturan (Rule) pada Rapidminer**4. KESIMPULAN**

Berdasarkan penelitian menentukan dapat diperoleh suatu model aturan yang dapat memperlihatkan bahwa faktor yang mempengaruhi Kepuasan Nasabah Terhadap Pelayanan Klaim di PT. Asuransi Central Asia pada *node* pertama adalah Faktor Assurance, *node* kedua Faktor Reliability, *node* ketiga Faktor Lama Nasabah. Untuk label Besar model aturan terbaik adalah Jika Assurance = baik, dan Reliability = baik maka hasilnya Puas { Puas = 16 dan Tidak Puas = 0} dengan jumlah atribut 16 Label 1 dan label Kecil model aturan terbaik adalah Jika Assurance = sangat baik maka hasilnya Puas {Puas = 15 dan Tidak Puas = 0} Label 9. Algoritma C4.5 dapat menghasilkan *rule* untuk menggambarkan proses yang terkait dengan faktor yang mempengaruhi Kepuasan Nasabah Terhadap Pelayanan Klaim di PT. Asuransi Central Asia dengan tanggapan dari nasabah yang Puas dan Tidak Puas.

REFERENCES

- [1] Akhmad Zulkifli. (2016). Metode C45 Untuk Mengklarifikasi Pelanggan Perusahaan Telekomunikasi Seluler Akhmad. *Revista Brasileira de Ergonomia*, 9(2), 10. <https://doi.org/10.5151/cidi2017-060>
- [2] Asmira. (2019). Penerapan Data Mining untuk Mengklasifikasi Pola Nasabah Menggunakan Algoritma C4,5 pada Bank BRI Unit Andounohu Kendari. *Jurnal Sistem Komputer Dan Sistem Informasi*, 1(1), 22–28.
- [3] Dimas Bayu Febriyanto. (2018). Implementasi Algoritma C4 . 5 Untuk Klasifikasi Tingkat Kepuasan Pembeli Online Shop. *Jurnal Riset Komputer*, 5(6), 569–575.
- [4] Faisal, S. (2019). Klasifikasi Data Minning Menggunakan Algoritma C4.5 Terhadap Kepuasan Pelanggan Sewa Kamera Cikarang. *Techno Xplore: Jurnal Ilmu Komputer Dan Teknologi Informasi*, 4(1), 1–8. <https://doi.org/10.36805/technoxplore.v4i1.541>
- [5] Gerardo, B. G., Saifullah, S., & Irawan, E. (2019). Teknik Data Mining Dalam Penilaian Pengajaran Guru Berdasarkan Indeks Kepuasan Siswa. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 508–514. <https://doi.org/10.30865/komik.v3i1.1634>
- [6] Jamhur, A. I. (2016). Penerapan Data Mining Untuk Menganalisa Jumlah Pelanggan Aktif Dengan Menggunakan Algoritma C4.5. *Majalah Ilmiah*, Vol. 23(No.2), 12–20.
- [7] Krismianto, U. (2015). SPSS (Statistical Package for the Social Sciens) Makalah Olah Data SPSS. In *Data SPSS (Issue Makalah Olah Data SPSS Kelembaban Udara 5 Kota Provinsi Jateng)*.
- [8] Lambertus, S. (2018). *Optimasi Kinerja Firewall Menggunakan Teknik Data Mining*.
- [9] Rani, L. N. (2016). Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit. *INOVTEK Polbeng - Seri Informatika*, 1(2), 126. <https://doi.org/10.35314/isi.v1i2.131>
- [10] Ropianto, M. (2016). Pemahaman Penggunaan Unified Modelling Language. *Jt-Ibsi*, 1(1), 43–50.
- [11] Sandrawira Anggraini, Sarjon Defit, G. W. N. (2018). Analisis Data Mining Penjualan Ban Menggunakan. *Jurnal Ilmu Teknik Elektro Komputer Dan Informatika (JITEKI)*, 4(2), 136–143.
- [12] Sembiring, M. A., Sibuea, M. F. L., & Sapta, A. (2018). Analisa Kinerja Algoritma C4.5 Dalam Memprediksi Hasil Belajar. *JSSR, February*, 73–79.
- [13] Setiyadi, D., Tetap, D., Eresha, S., Laboratorium, S., & Trisakti, U. (2011). Data Mining Dengan Metode Decision Tree Algoritma C4 . 5 Untuk Memprediksi Permintaan Jenis Produk Barang. *Jurnal Teknologi Informasi ESIT*, VI(02), 13–34.
- [14] Sitorus, L. (2015). *Algoritma Dan Pemrograman* (A. Pramesta (ed.); 1st ed.). ANDI.
- [15] Tanjung, Q. (2017). *Sistem Penunjang Keputusan Kelayakan Kredit Motor Menggunakan Metode Naive Bayes Pada NSC Finance Cikampek*.
- [16] Urva, G., & Siregar, H. F. (2015). Pemodelan UML E-Marketing Minyak Goreng. *JTSI*, 1(9), 92–101.