

IMPLEMENTASI DATA MINING UNTUK PENGENALAN KARAKTERISTIK TRANSAKSI CUSTOMER DENGAN MENGGUNAKAN ALGORITMA C4.5

Sarah Faradillah (12110837)

Mahasiswa Program Studi Teknik Informatika STMIK Budi Darma Medan
Jl. Sisingamangaraja No. 338 Sp. Limun Medan
www.stmik-budidarma.ac.id //Email: faradillahsarah@gmail.com

ABSTRAK

Pengenalan karakteristik transaksi customer dengan menggunakan data - data yang telah ada di masa lalu untuk memprediksi kebutuhan para customer dimasa yang akan datang sangat diperlukan oleh perusahaan. Agar menjadi acuan untuk perusahaan meningkatkan hasil produknya. Data mining dapat digunakan untuk melakukan analisis data dan menemukan pola-pola penting pada data. Data mining tersebut akan menjadi tolak ukur ataupun acuan untuk mengambil keputusan. Pengolahan data mining dapat dilakukan dengan Algoritma C4.5. Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Adapun perangkat lunak yang menerapkan Algoritma C4.5 ini adalah Tanagra 1.4.

Kata kunci : Data Mining, Karakteristik Transaksi, Algoritma C4.5, Tanagra 1.4

1. Pendahuluan

1.1. Latar Belakang Masalah

Sulitnya memprediksi kebutuhan pasar, merupakan kendala yang dihadapi perusahaan distribusi sehingga manajemen harus dapat mengambil keputusan yang tepat dan cepat guna memberikan pelayanan yang baik serta kepuasan kepada *customer* maupun menjaga konsistensi produk perusahaan di pasar. Keputusan yang diambil harus mempertimbangkan dengan baik berdasarkan data-data yang dimiliki, terutama yang berkaitan erat dengan transaksi *customer*. Pengenalan karakteristik transaksi *customer* dengan menggunakan data - data yang telah ada di masa lalu untuk memprediksi kebutuhan para *customer* dimasa yang akan datang sangat diperlukan oleh perusahaan. Agar menjadi acuan untuk perusahaan meningkatkan hasil produknya, dan untuk bagian *marketing* bisa meningkatkan penjualan melalui promosi penawaran dengan memberikan hal - hal yang baru pada produk dan layanan terhadap *customer*. Karena *customer* sebagian besar memiliki sifat mencari hal - hal baru daripada bertahan pada produk yang sudah mereka gunakan.

Dalam menganalisa hal tersebut di atas, salah satu solusi yang dapat digunakan adalah *data mining* mengenai kriteria - kriteria tertentu di dalam pengenalan karakteristik transaksi *customer*. Dengan semakin besarnya jumlah data dan kebutuhan akan analisis data yang akurat maka dibutuhkan metode analisis data yang tepat. Pengolahan *data mining* dapat dilakukan dengan beberapa teknik diantaranya adalah Metode *Rough Set*, *Fuzzy*, Algoritma *Apriori*, Algoritma C4.5 dan lain - lain.

Algoritma C4.5 merupakan algoritma yang

digunakan untuk membentuk pohon keputusan. Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Dengan masing-masing rangkaian pembagian, anggota himpunan hasil menjadi mirip satu dengan yang lain. Sebuah model pohon keputusan terdiri dari sekumpulan aturan untuk membagi sejumlah populasi yang heterogen menjadi lebih kecil, lebih *homogen* dengan memperhatikan pada variabel tujuannya.

1.2 Perumusan Masalah

Sesuai dengan uraian dalam pendahuluan di atas, yang menjadi permasalahannya adalah :

1. Bagaimana mengenali karakteristik transaksi *customer* PD. Cipta Sari Mandiri Tanjung Morawa menggunakan Algoritma C4.5 ?
2. Bagaimana menerapkan algoritma C4.5 dalam pengenalan karakteristik transaksi *customer* PD. Cipta Sari Mandiri Tanjung Morawa ?

1.3 Tujuan

Ada beberapa tujuan yang akan dicapai, yaitu sebagai berikut :

1. Untuk mengenali data karakteristik transaksi dari setiap *customer* PD. Cipta Sari Mandiri Tanjung Morawa menggunakan Algoritma C4.5.
2. Untuk menerapkan algoritma C4.5 dalam pengenalan karakteristik transaksi *customer* PD. Cipta Sari Mandiri Tanjung Morawa.

2. Landasan Teori

2.1 Implementasi

2.2

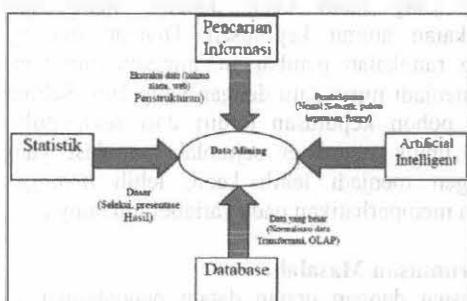
Menurut Nurdin Usman dalam bukunya Konteks Implementasi Berbasis Kurikulum mengemukakan Implementasi adalah bermuara pada aktivitas, aksi, tindakan, atau adanya mekanisme suatu sistem. Implementasi bukan sekedar aktivitas, tetapi suatu kegiatan yang terencana dan untuk mencapai tujuan kegiatan.

2.3 Data Mining

Menurut Larose *Data mining* adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang terselubung di dalam *database*. *Data mining* merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan yang terkait dari berbagai *database* besar.

Bidang Ilmu Data Mining

Data mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan *data mining* adalah kenyataan bahwa *data mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu.



Gambar 1. Bidang ilmu Data Mining

2.4 Pohon Keputusan

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Menurut Berry & Linoff Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Dengan masing-masing rangkaian pembagian, anggota himpunan hasil menjadi mirip satu dengan yang lain.

Sebuah model pohon keputusan terdiri dari sekumpulan aturan untuk membagi sejumlah populasi yang heterogen menjadi lebih kecil, lebih homogen dengan memperhatikan pada variabel tujuannya. Sebuah pohon keputusan mungkin dibangun dengan seksama secara manual, atau dapat tumbuh secara otomatis dengan menerapkan salah satu atau beberapa algoritma pohon keputusan

untuk memodelkan himpunan data yang belum terklasifikasi.

Variabel tujuan biasanya dikelompokkan dengan pasti dan model pohon keputusan lebih mengarah pada perhitungan probabilitas dari masing-masing *record* terhadap kategori-kategori tersebut, atau untuk mengklasifikasi *record* dengan mengelompokkannya dalam satu kelas. Pohon keputusan juga dapat digunakan untuk mengestimasi nilai dari variabel *continues*, meskipun ada beberapa teknik yang lebih sesuai untuk kasus ini.

2.5 Algoritma C4.5

Algoritma C 4.5 adalah salah satu metode untuk membuat *decision tree* berdasarkan *training data* yang telah disediakan. Algoritma C4.5 merupakan pengembangan dari ID3. Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi *missing value*, bisa mengatasi *continues data*, dan *pruning*.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. prose

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

Ulan gi

s untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *Gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *Gain* digunakan rumus seperti tertera dalam persamaan berikut :

... (1)

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

Keterangan :
S : Himpunan

Kasus

A : Atribut

n : Jumlah partisi atribut A

|S_i| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

Sementara itu perhitungan nilai *entropy* dapat dilihat pada persamaan berikut :

..... (2)

Keterangan :

S : Himpunan Kasus

n : Jumlah partisi S

p_i : Proporsi dari S_i terhadap S

2.6 Karakteristik Pelanggan

- Menurut Griffin mengungkapkan bahwa pelanggan memiliki karakteristik sebagai berikut :
1. Melakukan pembelian secara teratur.
 2. Membeli di luar lini produk atau jasa.
 3. Merekomendasikan kepada orang lain.
 4. Menunjukkan kekebalan terhadap daya tarik produk sejenis dari pesaing.

2.7 Tanagra

Tanagra merupakan salah satu *software data mining* yang didalamnya disediakan beberapa metode *data mining* mulai dari mengekplorasi analisis data, pembelajaran statistik, pembelajaran mesin, dan *database*.

Tidak seperti *software data mining* kebanyakan, tanagra merupakan suatu *software* berbasis *open source* di mana semua orang dapat mengakses *source codenya*, dan menambahkan algoritma mereka sendiri, sejauh dia setuju dan menyesuaikan dengan lisensi pendistribusian *softwarenya*.

3. Analisa Dan Pembahasan

3.1 Analisa Data Penjualan

Analisa data penjualan dilakukan setelah data terkumpul dan sesuai dengan kebutuhan sistem ini. Oleh sebab itu, untuk menghasilkan kesimpulan berdasarkan pohon keputusan pada analisa data ini, diperlukan data produk yang telah dibeli *customer*. Analisa data tersebut dilakukan berdasarkan teknik algoritma C4.5 dengan beberapa langkah-langkah yang sudah ditentukan.

Tabel 1. Produk yang dibeli customer

NO	Nama Customer	Jenis Produk Yang Dibeli					Status Customer	
		TB. Primeo	TB. Saguku	MSG PM	MSG RMM	MSG RSK		
1	Citra	Beli	Beli	Tidak Beli	Tidak Beli	Beli	Tidak Beli	Tetap
2	Nakidi	Beli	Tidak Beli	Beli	Tidak Beli	Beli	Tidak Beli	Tetap
3	Riska	Beli	Tidak Beli	Beli	Tidak Beli	Tidak Beli	Beli	Tetap
4	Hasmid	Beli	Tidak Beli	Tidak Beli	Beli	Beli	Tidak Beli	Baru
5	Azni	Tidak Beli	Beli	Beli	Beli	Beli	Beli	Tetap
6	Salamun	Tidak Beli	Beli	Beli	Tidak Beli	Beli	Beli	Tetap
7	Kemar	Beli	Tidak Beli	Tidak Beli	Tidak Beli	Beli	Beli	Baru
8	Dewi	Beli	Beli	Beli	Beli	Beli	Tidak Beli	Tetap
9	Sukudi	Tidak Beli	Tidak Beli	Beli	Beli	Tidak Beli	Beli	Tetap
10	Ramah	Beli	Tidak Beli	Tidak Beli	Beli	Tidak Beli	Beli	Baru
11	Wahyu	Beli	Beli	Beli	Beli	Tidak Beli	Beli	Tetap
12	Rahayu	Beli	Beli	Beli	Tidak Beli	Beli	Beli	Tetap
13	Ims	Tidak Beli	Tidak Beli	Beli	Beli	Tidak Beli	Tidak Beli	Tetap
14	Ajura	Beli	Tidak Beli	Beli	Tidak Beli	Tidak Beli	Beli	Tetap
15	Aci	Tidak Beli	Beli	Beli	Beli	Beli	Beli	Tetap
16	Yenny	Beli	Tidak Beli	Beli	Tidak Beli	Tidak Beli	Beli	Baru
17	Aya	Beli	Beli	Tidak Beli	Beli	Beli	Beli	Tetap
18	Aho	Tidak Beli	Tidak Beli	Tidak Beli	Beli	Beli	Beli	Baru
19	Pengajuan	Beli	Beli	Beli	Beli	Tidak Beli	Beli	Tetap
20	Foko	Tidak Beli	Tidak Beli	Tidak Beli	Tidak Beli	Tidak Beli	Beli	Baru
21	Seri	Beli	Tidak Beli	Beli	Beli	Beli	Beli	Tetap
22	Eko	Beli	Beli	Tidak Beli	Beli	Tidak Beli	Beli	Tetap
23	Ika	Beli	Beli	Beli	Beli	Tidak Beli	Beli	Tetap
24	Ha	Tidak Beli	Tidak Beli	Beli	Tidak Beli	Beli	Tidak Beli	Baru

3.1.1 Menghitung Jumlah Penjualan Berdasarkan Semua Entropy (Produk)

Menghitung jumlah penjualan, antara lain jumlah penjualan untuk keputusan *customer* tetap, jumlah penjualan untuk keputusan *customer* baru, dan *entropy* dari semua penjualan dan penjualan yang dibagi berdasarkan atribut TB. Primeo, TB. Saguku, MSG PM, MSG RMM, MSG RSK, CKM. Setelah itu melakukan perhitungan *gain* dari setiap atribut.

Baris total kolom *entropy* pada Tabel 2 dihitung dengan rumus 2 sebagai berikut :

$$\text{Entropy(Total)} = \left(-\frac{17}{24} * \log_2 \left(\frac{17}{24} \right) \right) + \left(-\frac{7}{24} * \log_2 \left(\frac{7}{24} \right) \right) \dots$$

(3)

Sementara itu, nilai *entropy* pada baris TB. Primeo beli dihitung dengan menggunakan rumus 2 berikut :

$$\text{Entropy(TB.Primeo,beli)} = \left(-\frac{12}{16} * \log_2 \left(\frac{12}{16} \right) \right) + \left(-\frac{4}{16} * \log_2 \left(\frac{4}{16} \right) \right) = 0,81$$

Nilai *entropy* pada baris TB. Primeo tidak beli dihitung dengan menggunakan rumus 2 berikut:

$$\text{Entropy(TB.Primeo,tidak beli)} = \left(-\frac{5}{8} * \log_2 \left(\frac{5}{8} \right) \right) + \left(-\frac{3}{8} * \log_2 \left(\frac{3}{8} \right) \right) = 0,95$$

Sementara itu, nilai *gain* pada baris TB. Primeo dihitung dengan menggunakan rumus 1 berikut :

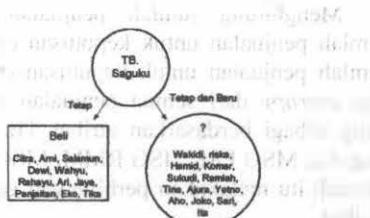
$$\text{Gain(total,TB.Primeo)} = 0,87 - \left(\left(\frac{16}{24} * 0,81 \right) + \left(\frac{8}{24} * 0,95 \right) \right) = 0,01$$

Tabel 2. Perhitungan Node 1

Node			Jumlah (S)	Status Customer		Entropy	Gain
				Temp (S1)	Baru (S2)		
1	Total		24	17	7	0,87	
	TB. Primeo						0,01
	Beli		16	12	4	0,81	
	Tidak Beli		8	5	3	0,95	
	TB. Saguku						0,33
	Beli		11	11	0	0	
	Tidak Beli		13	6	7	1	
	MSG PM						0,15
	Beli		15	13	2	0,57	
	Tidak Beli		9	4	5	0,99	
	MSG RMM						0,20
	Beli		13	12	1	0,39	
	Tidak Beli		11	5	6	0,99	
	MSG RSK						0,02
	Beli		14	9	5	0,94	
	Tidak Beli		10	8	2	0,72	
	CKM						0,05
	Beli		15	12	3	0,72	
	Tidak Beli		9	5	4	0,99	

Dari hasil Tabel 4.2 dapat diketahui bahwa atribut dengan *gain* tertinggi adalah produk TB. Saguku, yaitu sebesar 0,33. Dengan demikian, TB. Saguku dapat menjadi *Node* akar. Ada dua atribut dari produk TB. Saguku, yaitu beli dan tidak beli. Dari kedua nilai atribut tersebut, nilai atribut beli sudah mengklasifikasikan penjualan menjadi satu, yaitu keputusan *customer* tetap, sehingga tidak perlu dilakukan perhitungan lebih lanjut, tetapi untuk nilai atribut tidak beli masih perlu dilakukan perhitungan lagi.

Dari hasil tersebut dapat digambarkan pohon keputusannya sementara tampak seperti Gambar 2 dibawah ini :



Gambar 2. Pohon Keputusan Hasil Perhitungan Node 1

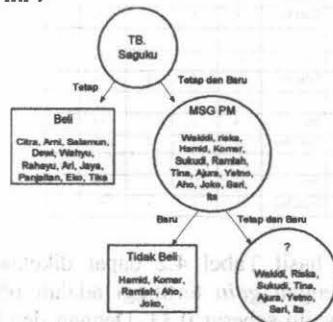
3.1.2 Menghitung Jumlah Penjualan Berdasarkan Produk TB. Primeo, MSG RMM, MSG PM, MSG RSK Dan CKM

Menghitung jumlah penjualan, antara lain jumlah penjualan untuk keputusan *customer* tetap, jumlah penjualan untuk keputusan *customer* baru dan entropy dari semua penjualan. Dan penjualan yang dibagi berdasarkan atribut TB. Primeo, MSG. RMM, MSG PM, MSG RSK dan CKM. Setelah itu melakukan perhitungan *gain* dari tiap-tiap atribut. Hasil perhitungan dapat dilihat pada Tabel 3.

Tabel 3. Perhitungan Node 2

Node	Jumlah (S)	Status Customer	Entropy	Gain
2 Total	13	Tetap (S1)	7	1
TB. Sagaku - Tidak Beli				0,01
TB. Primeo				0,01
Beli	3	4	4	1
Tidak Beli	5	2	3	0,97
MSG RMM				0,14
Beli	4	3	1	0,81
Tidak Beli	9	3	6	0,92
MSG PM				0,49
Beli	6	6	2	0,81
Tidak Beli	5	0	5	0
MSG RSK				0,11
Beli	7	2	5	0,86
Tidak Beli	6	4	2	0,92
CKM				0,04
Beli	7	4	3	0,99
Tidak Beli	6	2	4	0,95

Dari hasil tersebut dapat digambarkan pohon keputusan sementaranya tampak seperti Gambar 3 dibawah ini :



Gambar 3. Pohon Keputusan Hasil Perhitungan Node 2

3.1.3 Menghitung Jumlah Penjualan Berdasarkan Produk TB. Primeo, MSG RMM, MSG RSK dan CKM

Menghitung jumlah penjualan, antara lain jumlah penjualan untuk keputusan *customer* tetap, jumlah penjualan untuk keputusan *customer* baru, dan entropy dari semua penjualan dan penjualan yang dibagi berdasarkan atribut TB. Primeo, MSG RMM, MSG RSK dan CKM. Setelah itu

menghitung perhitungan *gain* dari tiap-tiap atribut. Hasil perhitungan dapat dilihat pada tabel dibawah ini.

Tabel 4. Perhitungan Node 3

Node	Jumlah (S)	Status Customer	Entropy	Gain
3 Total	8	Tetap (S1)	2	0,81
TB. Sagaku - Tidak Beli MSG PM -Beli				0,02
TB. Primeo				0,37
Beli	5	4	1	0,72
Tidak Beli	3	2	1	0,92
MSG RMM				-0,03
Beli	3	3	0	0
Tidak Beli	5	3	2	0,97
MSG RSK				-0,03
Beli	5	2	1	0,92
Tidak Beli	3	4	1	0,72
CKM				0,44
Beli	4	4	0	0
Tidak Beli	4	2	2	1

Dari hasil tersebut dapat digambarkan pohon keputusan sementaranya tampak seperti Gambar 4 dibawah ini :



Gambar 4. Pohon Keputusan Hasil Perhitungan Node 3

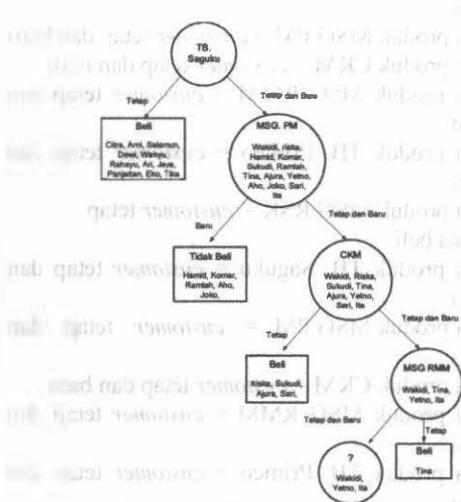
3.1.4 Menghitung Jumlah Penjualan Berdasarkan Produk TB. Primeo, MSG RMM dan MSG RSK

Menghitung jumlah penjualan, jumlah penjualan untuk keputusan *customer* tetap, jumlah penjualan untuk keputusan *customer* baru, dan entropy dari semua penjualan dan penjualan yang dibagi berdasarkan atribut TB. Primeo, MSG RMM dan MSG RSK Setelah itu melakukan perhitungan *gain* dari tiap-tiap atribut. Hasil perhitungan dapat dilihat pada Tabel 5 dibawah ini.

Tabel 5. Perhitungan Node 4

Node	Jumlah (S)	Status Customer	Entropy	Gain
4 Total	4	Tetap (S1)	2	1
TB. Sagaku - Tidak Beli MSG PM -Beli CKM -Tidak Beli				0
TB. Primeo				0,78
Beli	2	1	1	1
Tidak Beli	2	1	1	1
MSG RMM				0,78
Beli	1	1	0	0
Tidak Beli	3	1	2	0,92
MSG RSK				0
Beli	2	1	1	1
Tidak Beli	2	1	1	1

Dari hasil tersebut dapat digambarkan pohon keputusan sementaranya tampak seperti Gambar 5 dibawah ini :



Gambar 5. Pobon Keputusan Hasil Perhitungan Node 4

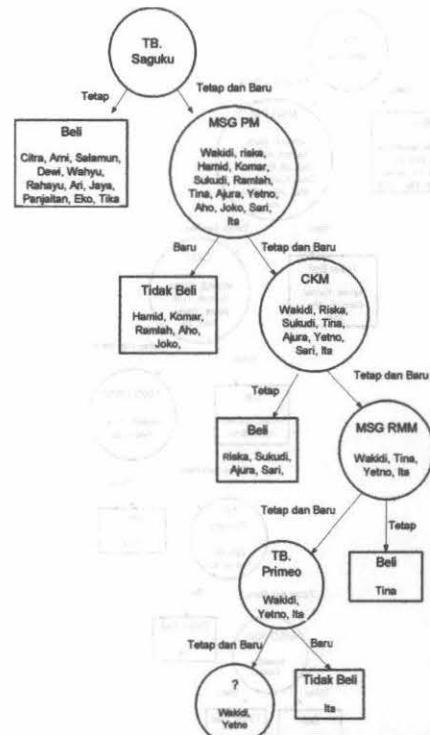
3.1.5 Menghitung Jumlah Penjualan Berdasarkan Produk TB, Primeo Dan MSG RSK

Menghitung jumlah penjualan, jumlah penjualan untuk keputusan *customer* tetap, jumlah penjualan untuk keputusan *customer* baru, dan *entropy* dari semua penjualan dan penjualan yang dibagi berdasarkan atribut TB, Primeo dan MSG RSK. Setelah itu melakukan perhitungan *gain* dari tiap-tiap atribut. Hasil perhitungan dapat dilihat pada Tabel 6 dibawah ini.

Tabel 6. Perhitungan Node 5

Node			Jumlah (S)	Status Customer		Entropy	Gain
				Tetap (S1)	Baru (S2)		
5	Total		3	1	2	0,92	
	TB, Saguku						
	-Tidak Beli						
	MSG P34						
	-Belii						
	CQM						
	-Tidak Beli						
	MSG RMM						
	-Tidak Beli						
	TB. Primeo						0,33
	Beli		2	1	1	1	
	Tidak Beli		1	0	1	0	
	MSG RSK						
	Beli		2	1	1	1	0,25
	Tidak Beli		1		1	0	

Dari hasil tersebut dapat digambarkan pohon keputusan sementaranya tampak seperti Gambar 6 dibawah ini :



Gambar 6. Pobon Keputusan Hasil Perbitungan Node 5

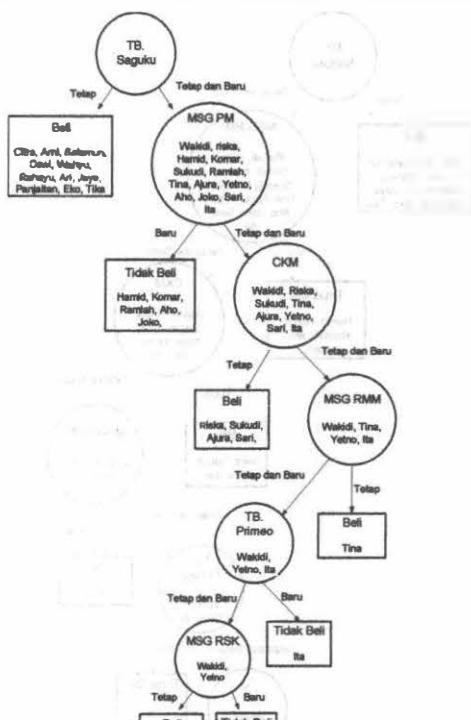
3.1.6 Mengbitung Jumlah Penjualan Berdasarkan Produk MSG RSK

Menghitung jumlah penjualan, jumlah penjualan untuk keputusan *customer* tetap, jumlah penjualan untuk keputusan *customer* baru, dan *entropy* dari penjualan berdasarkan atribut MSG RSK. Setelah itu melakukan perhitungan *gain* dari tiap-tiap atribut. Hasil perhitungan dapat dilihat pada Tabel 7 dibawah ini.

Tabel 7. Perhitungan Node 6

Node	Pertanyaan	Jumlah (S)	Status Customer		Entropy	Gain
			Tetap (\\$1)	Baru (\\$2)		
6	Total	2	1	1	0	0
	TB. Saguna					
	-Tidak Beli					
	MSG PMI					
	-Beli					
	CSK					
	-Tidak Beli					
	MSG RSK					
	-Beli	1	1	0	0	1
	Tidak Beli	1	0	1	0	0

Dari hasil tersebut dapat digambarkan pohon keputusan akhir seperti Gambar 7 dibawah ini :



Gambar 7. Pohon Keputusan Hasil Perhitungan Node 6

Berdasarkan Gambar 7 dapat diambil daftar aturan dari pohon keputusan yaitu :

1. Jika produk TB. Saguku = *customer* tetap
Maka Beli
2. Jika produk TB. Saguku = *customer* tetap dan baru
Jika produk MSG PM = *customer* baru
Maka Tidak beli
3. Jika produk TB. Saguku = *customer* tetap dan baru
Jika produk MSG PM = *customer* tetap dan baru
Jika produk CKM = *customer* tetap
Maka beli
4. Jika produk TB. Saguku = *customer* tetap dan baru
Jika produk MSG PM = *customer* tetap dan baru
Jika produk CKM = *customer* tetap dan baru
Jika produk MSG RMM = *customer* Tetap
Maka beli
5. Jika produk TB. Saguku = *customer* tetap dan baru
Jika produk MSG PM = *customer* tetap dan baru
Jika produk CKM = *customer* tetap dan baru
Jika produk MSG RMM = *customer* tetap dan baru
Jika produk TB. Primeo = *customer* baru
Maka tidak beli
6. Jika produk TB. Saguku = *customer* tetap dan

baru

Jika produk MSG PM = *customer* tetap dan baru

Jika produk CKM = *customer* tetap dan baru

Jika produk MSG RMM = *customer* tetap dan baru

Jika produk TB. Primeo = *customer* tetap dan baru

Jika produk MSG RSK= *customer* tetap

Maka beli

7. Jika produk TB. Saguku = *customer* tetap dan baru

Jika produk MSG PM = *customer* tetap dan baru

Jika produk CKM = *customer* tetap dan baru

Jika produk MSG RMM = *customer* tetap dan baru

Jika produk TB. Primeo = *customer* tetap dan baru

Jika produk MSG RSK= *customer* baru

Maka tidak beli

Berdasarkan daftar aturan diatas dapat diambil kesimpulan yaitu :

1. Produk Tb. Saguku dibeli oleh *customer* tetap, maka pihak gudang bisa menyiapkan stock produk Tb. Saguku untuk kedepannya.
2. Produk MSG PM tidak dibeli oleh *customer* baru, maka pihak *marketing* harus meningkatkan penawaran dengan memberikan hal-hal yang baru pada produk dan layanan terhadap *customer* agar bisa menarik perhatian para *customer* untuk menggunakan produk MSG PM ini.
3. Produk CKM dibeli oleh *customer* tetap, maka pihak gudang bisa menyiapkan stock produk CKM untuk kedepannya.
4. Produk MSG RMM dibeli oleh *customer* tetap, maka pihak gudang bisa menyiapkan stock produk MSG RMM untuk kedepannya.
5. Produk TB. Primeo tidak dibeli oleh *customer* baru, maka pihak *marketing* harus meningkatkan penawaran dengan memberikan hal-hal yang baru pada produk dan layanan terhadap *customer* agar bisa menarik perhatian para *customer* untuk menggunakan produk TB. Primeo ini.
6. Produk MSG RSK dibeli oleh *customer* tetap, maka pihak gudang bisa menyiapkan stock produk MSG RSK untuk kedepannya.
7. Produk MSG RSK tidak dibeli oleh *customer* baru, maka pihak *marketing* harus meningkatkan penawaran dengan memberikan hal-hal yang baru pada produk dan layanan terhadap *customer* agar bisa menarik perhatian para *customer* untuk menggunakan produk MSG RSK ini.

4. Implementasi

4.2 Pengujian

Pengujian perangkat lunak merupakan suatu

investigasi yang dilakukan untuk mendapatkan informasi mengenai kualitas dari produk atau layanan yang sedang diuji. Pengujian perangkat lunak juga memberikan pandangan mengenai perangkat lunak secara obyektif dan independen, yang bermanfaat dalam operasional bisnis untuk memahami tingkat resiko pada implementasinya. Teknik-teknik pengujian mencakup, namun tidak terbatas pada proses mengeksekusi suatu bagian program atau keseluruhan aplikasi dengan tujuan untuk menemukan kesalahan atau cacat lainnya. Adapun pengujian dalam perangkat lunak Tanagra 1.4 ini dapat dilakukan dalam perangkat komputer.

4.2.1 Input file

Setelah semua kebutuhan *hardware* dan *software* dilakukan, selanjutnya melakukan penginputan nilai data *customer* ke dalam Microsoft Excel.

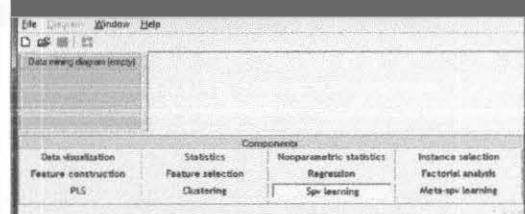
	A	B	C	D	E	F	G
	TB_Promo	TB_Begaku	MSG_PM	MSG_PBM	MSG_PSK	class_customer	
1	0.99	0.99	0.99	0.99	0.99	Baru	
2	0.81	1	0.97	0.99	0.99	Teng	
3	0.81	1	0.97	0.99	0.99	Teng	
4	0.81	1	0.97	0.99	0.99	Teng	
5	0.81	1	0.99	0.99	0.99	Baru	
6	0.85	0	0.99	0.99	0.99	Teng	
7	0.85	0	0.99	0.99	0.99	Teng	
8	0.81	1	0.99	0.99	0.99	Baru	
9	0.81	0	0.97	0.99	0.99	Teng	
10	0.85	1	0.97	0.99	0.99	Teng	
11	0.81	1	0.99	0.99	0.99	Baru	
12	0.81	0	0.97	0.99	0.99	Teng	
13	0.81	0	0.97	0.99	0.99	Teng	
14	0.85	1	0.97	0.99	0.99	Baru	
15	0.81	1	0.97	0.99	0.99	Teng	
16	0.81	1	0.99	0.99	0.99	Baru	
17	0.81	1	0.97	0.99	0.99	Baru	
18	0.81	0	0.99	0.99	0.99	Teng	
19	0.85	1	0.99	0.99	0.99	Baru	
20	0.81	0	0.97	0.99	0.99	Teng	
21	0.85	1	0.99	0.99	0.99	Baru	
22	0.81	1	0.97	0.99	0.99	Teng	
23	0.81	0	0.99	0.99	0.99	Teng	
24	0.81	0	0.97	0.99	0.99	Teng	
25	0.85	1	0.97	0.99	0.99	Baru	

Gambar 8. Data Customer Dalam Hasil Perhitungan Entropy

Setelah nilai data *customer* di *input* ke dalam Microsoft Excel, lalu disimpan dengan format *Text* (*Tab delimited*).

4.2.2 Tampilan Tanagra 1.4

Adapun tampilan perangkat lunak Tanagra 1.4 yaitu pada Gambar 9.

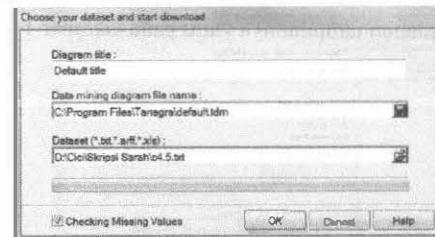


Gambar 9. Tampilan Tanagra 1.4

4.2.3 Tampilan Download

Sebelum data diolah maka data harus di *download* terlebih dahulu atau tepatnya pilih *file*. Adapun tampilannya yaitu pada Gambar 10.

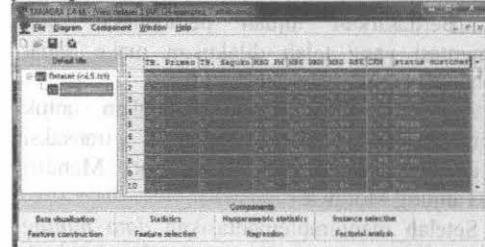
Implementasi Data Mining Untuk Pengenalan Karakteristik Transaksi Customer Dengan Menggunakan Algoritma C4.5. Sarah Faradillah



Gambar 10. Tampilan Download

4.2.4 Tampilan View DataSet

Untuk melihat tampilan data yang telah dipilih, maka digunakan *viewdataset*. Adapun tampilannya yaitu pada Gambar 11.



Gambar 11. Tampilan View dataset

4.2.5 Tampilan Define Attribute Statuses

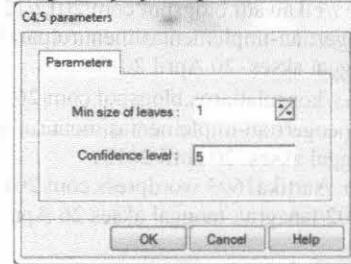
Define Attribute Statuses digunakan untuk menginput atribut dan menginput target. Adapun tampilannya yaitu pada Gambar 12.



Gambar 12. Tampilan Define Attribute Statuses

4.2.6 Tampilan C4.5 Parameters

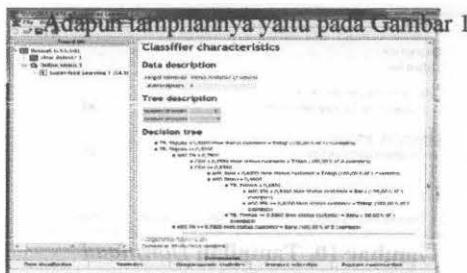
C4.5 Parameters digunakan untuk menginput minimal *size of leaves* dan *confidence level*. Adapun tampilannya yaitu pada Gambar 13.



Gambar 13. Tampilan C4.5 parameters

4.2.7 Tampilan Supervised Learning

Supervised Learning merupakan tahap akhir untuk mengetahui *decision tree* (pohon keputusan).



Gambar 14. Tampilan *Supervised Learning*

5. Kesimpulan

Berdasarkan tujuan pembahasan dan implementasi yang telah dilakukan maka dapat diperoleh beberapa kesimpulan yaitu :

1. Algoritma C4.5 dapat digunakan untuk mengenali pola karakteristik transaksi *customer* pada PD. Cipta Sari Mandiri Tanjung Morawa.
 2. Setelah mengimplementasikan *data mining* dengan algoritma C4.5 maka karakteristik transaksi *customer* sudah dapat dikenali sehingga pihak PD. Cipta Sari Mandiri Tanjung Morawa bisa mengambil keputusan yang berkaitan dengan menyiapkan atau menyediakan *stock* produk yang digunakan untuk masa yang akan datang. Dan untuk bagian pemasaran harus meningkatkan penawaran-penawaran dengan memberikan hal-hal yang baru pada produk untuk menarik perhatian para *customer* menggunakan produk yang ditawarkan.

DAFTAR PUSTAKA

- [1] Kusrini, dan Luthfi, T. L, "Algoritma Data Mining", Edisi Pertama, Andi Offset, Yogyakarta, 2009
 - [2] Suyanto, "Artificial Intelligence", Edisi Revisi, Informatika Bandung, Bandung, 2011.
 - [3] <http://blog.unitomo.ac.id/lambang>, tanggal akses 20 April 2013
 - [4] <http://elkawaqi.blogspot.com/2012/12/pengertian-implementasimenumerutpara.html>, tanggal akses 20 April 2013
 - [5] <http://konsulatlaros.blogspot.com/2012/10/pengertian-implementasimenumerut.html>, tanggal akses 20 april 2013)
 - [6] <http://sartikal603.wordpress.com/2011/11/02/tanagra/>, tanggal akses 26 April 2011