

Prediksi Loyalitas Pelanggan Pada *Fast Moving Consumer Goods* Menggunakan Klasifikasi Metode C4.5

Predicting Customer Loyalty On Fast Moving Consumer Goods Using C4.5 Classification Method

Laila Isyriyah^{1*}
Imam Baihaqi²
Febry Eka Purwiantono³

^{1,2}Informatika, Sekolah Tinggi Informatika & Komputer Indonesia, Malang, Indonesia
³D3 Sistem Informasi, Sekolah Tinggi Informatika & Komputer Indonesia, Malang, Indonesia
¹laila@stiki.ac.id, ²imam.baihaqi2396@gmail.com, ³febry@stiki.ac.id

Penulis

Korespondensi:

Laila Isyriyah
laila@stiki.ac.id

Riwayat Artikel:

Diterima : 14 Desember 2023
Direview : 19 Desember 2023
Disetujui : 3 Januari 2024
Terbit : 5 Januari 2024

Abstrak

Mie instan merupakan salah satu produk makanan dari industri *Fast Moving Customer Goods* (FMCG) yang mana industri yang cukup besar di Indonesia. Namun, persaingan tidak bisa dihindari. Sehingga untuk memenangkan persaingan dari perusahaan lainnya, majamen perusahaan dituntut untuk menentukan strategi mempertahankan loyalitas konsumen. Oleh karena itu, tujuan penelitian ini adalah membuat suatu aplikasi untuk melakukan prediksi terhadap loyalitas pelanggan dan menentukan atribut yang berpengaruh dengan menerapkan Data Mining Klasifikasi dalam bentuk pohon keputusan. Metode aplikasi yang digunakan dalam Klasifikasi untuk prediksi adalah metode C4.5. Dalam algoritma C4.5 dilakukan perhitungan *entropy* dan *information gain* dimana atribut loyalitas pelanggan sebagai atribut tujuan (*class*), sedangkan harga, kemasan, citarasa, variasi, iklan, distribusi, dan kualitas sebagai atribut sumber untuk memperoleh node akar dan node lainnya. Hasil dari penelitian menunjukkan bahwa aplikasi dengan menggunakan metode C4.5 menghasilkan akurasi 97.5%, jadi metode C4.5 dapat digunakan untuk membantu manajemen perusahaan mie instan dalam rangka menentukan strategi mempertahankan loyalitas konsumen.

Kata Kunci: Prediksi; C4.5; Loyalitas Pelanggan; Data Mining; Klasifikasi

Abstract

Instant noodles are one of the food products from the Fast Moving Customer Goods (FMCG) industry which is a fairly large industry in Indonesia. However, competition is inevitable. So to win the competition from other companies, company management is required to determine a strategi to maintain customer loyalty. Therefore, the purpose of this study is to create an application to predict customer loyalty and determine the influential attributes by applying Data Mining Classification in the form of a decision tree. The application method used in Classification for prediction is the C4.5 method. In the C4.5 algorithm, entropy and information gain are calculated where customer loyalty is the attribute of destination (class), while price, packaging, taste, variety, advertising, distribution, and quality are the source attributes to obtain the root node and other nodes. The results of the study show that the application using the C4.5 method produces an

accuracy of 95.5%, so the C4.5 method can be used to assist the management of instant noodle companies in order to determine strategies to maintain consumer loyalty.

Keyword : Prediction; C4.5; Customer Loyalty; Data Mining; Classification.

1. Pendahuluan

Jasa penjualan mie instan membutuhkan adanya rancangan sistem yang akan diberikan kepada pelanggan untuk tetap mempertahankan loyalitas pelanggan pada perusahaan tersebut, dengan perkembangan teknologi persaingan antara penyedia jasa penjualan mie instan semakin ketat dengan adanya E-commerce yang menyediakan berbagai jasa penjualan mie instan. Perusahaan harus mampu mempertahankan sistem pasar yang menjadi sasaran dengan meningkatkan kualitas dari pelayanannya. Pelayanan yang berkualitas akan memberikan kepuasan bagi para pelanggan terhadap produk mie instan. Pelanggan yang telah mencapai kepuasan dalam pelayanan perusahaan akan mudah sekali mengajak teman, saudara atau orang lain dengan pengalaman mereka dalam menggunakan perusahaan mie instan. Pelanggan yang tidak puas akan menyebabkan pelanggan tersebut beralih ke perusahaan mie instan yang lain, kemungkinan besar akan bercerita kepada orang lain atas pelayanan yang dia dapatkan. Kepuasan pelanggan merupakan kelebihan dari yang diharapkan perusahaan mie instan untuk memenuhi keinginan dari pelanggan [1].

Perusahaan pangan, terutama produksi *Fast Moving Consumer Goods* (FMCG), sukses membantu 30% sampai dengan tahun 2023 akan produksi Produk Domestik Bruto (PDB) [1]. Di Indonesia, *Fast Moving Customer Goods* (FMCG) ialah salah satu perusahaan utama. Sektor FMCG adalah sektor yang melayani permintaan konsumen yang perlu bergerak dengan sangat cepat atau barang yang cepat habis. Produk FMCG memiliki umur yang relatif pendek karena sifatnya yang mudah rusak. FMCG juga diklasifikasikan secara terpisah, termasuk produksi obat-obatan (farmasi), produk elektronik, produksi makanan dan minuman kemasan [2].

Menurut beberapa data, industri FMCG memiliki pasar yang cukup besar di Indonesia. Sehingga untuk memenangkan persaingan yang dihadapi perusahaan, manajemen perusahaan harus lebih berhati-hati dalam menentukan strategi bersaingnya. Manajemen suatu perusahaan harus mampu mengembangkan dan menerapkan teknik pemasaran yang dapat mewujudkan, menjaga dan menumbuhkan loyalitas konsumen dengan demikian dapat menghasilkan loyalitas konsumen yang kuat atas barang yang dipasok. Setiap pemasar menawarkan berbagai manfaat dan daya tarik produk untuk memuaskan pelanggan. Produk makanan dan minuman, misalnya, merupakan salah satu jenis produk FMCG yang ada. Seiring berjalannya waktu, pasar makanan instan, khususnya mie instan, telah tumbuh dan berkembang dengan menyediakan berbagai macam pilihan.

Mie instan telah menjadi makanan pokok masyarakat Indonesia. Ada kecenderungan menganggapnya berbahaya bagi kesehatan, tapi kebiasaan makan mie instan sepertinya sulit dihilangkan. Alhasil, penjualan mie instan tidak akan menurun [3]. Menurut Asosiasi Mie Instan Dunia (WINA), per 13 Mei 2022, konsumsi mie instan di Indonesia sebesar 13,27 miliar porsi pada 2021, Jumlah tersebut naik 5,05% dibandingkan pada tahun sebelumnya yang sebanyak 12,64 miliar porsi. Posisi Indonesia berada tepat di bawah China/Hong Kong dengan konsumsi mie instan sebesar 43,99 miliar porsi pada tahun lalu Artikel ini telah tayang di DataIndonesia.id dengan judul "Konsumsi Mie Instan Indonesia Terbesar Kedua di Dunia pada 2021".

Jurnal penelitian sebelumnya dipakai untuk prediksi loyalitas pelanggan data seluler dimaksudkan untuk menerapkan versi klasifikasi pohon keputusan untuk memprediksi loyalitas pelanggan dan melihat variable yang paling berpengaruh. Tujuan dalam ulasan ini yaitu pelanggan dari data seluler. Data yang digunakan adalah data utama dari kuisioner

disebarluaskan yang berupa pertanyaan. Teknik yang dipakai untuk prediksi menerapkan metode C4.5. Pada algoritma C4.5 dilakukan perhitungan untuk mencari entropy dan information gain dimana atribut loyalitas dari pelanggan data seluler sebagai atribut tujuan sedangkan atribut lainnya sebagai atribut sumber untuk menentukan node akar dan node lainnya. Berdasarkan hasil klasifikasi menggunakan metode algoritma C4.5 menghasilkan akurasi mencapai 97.5%, yang memastikan bahwa algoritma C4.5 cocok digunakan untuk mengukur tingkat loyalitas pelanggan data seluler [4].

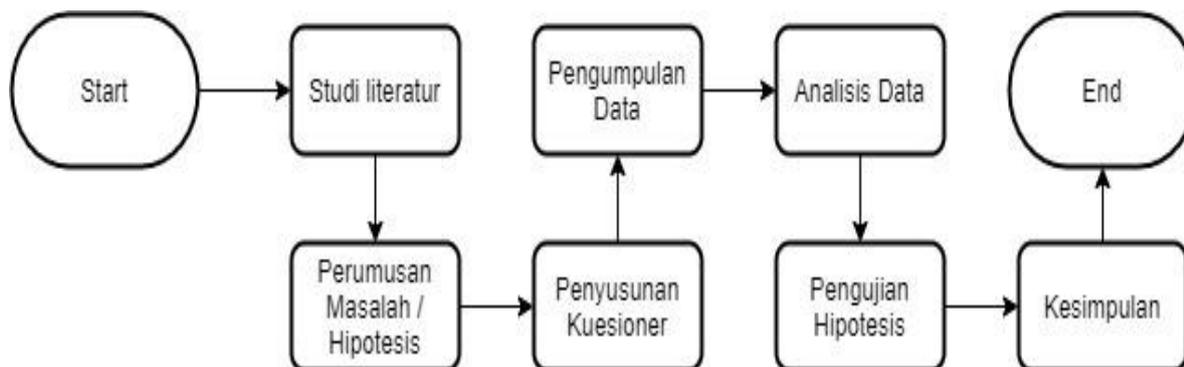
Sebab itu, tujuan dari penelitian ini dimaksudkan dapat membantu perusahaan mie instan untuk memprediksi loyalitas konsumen terhadap produk yang ditawarkan dengan cara survey kepada para konsumen mie instan dengan menyebarkan kuisioner berupa pertanyaan. Untuk itu peneliti menggunakan *data mining* dengan menggunakan metode C4.5.

Data mining merupakan sebutan yang dipakai pada proses menciptakan pengetahuan dalam database. Metode mengekstraksi dan mengenali informasi yang berguna dan mengumpulkan pengetahuan dari database besar memerlukan metode *statistical*, *mathematical*, *artificial intelligence*, dan *machine learning* dikenal sebagai Data Mining [5] [6].

Oleh karena itu, perlu adanya aplikasi desktop yang memungkinkan manager perusahaan lebih mudah memprediksi variabel yang paling berpengaruh dengan menggunakan metode C4.5. Metode C4.5 menurut Ibnu Fatchur Rohman (2015) ialah algoritma pembentuk sebuah *Decision Tree* (Pohon Keputusan) dan baik untuk memprediksi kepuasan pelanggan pada suatu perusahaan [7]. Pohon keputusan adalah salah satu dari teknissistem prediksi dan klasifikasi. Pohon keputusan membantu dalam eksplorasi data dan penemuan hubungan potensial antara sejumlah variabel input dan variabel tujuan. Teknik *decision tree* melibatkan mengonversikan bentuk data dalam bentuk versi *tree*, mengubah versi *tree* menjadi *rules*, dan memudahkan *rules*. Algoritma C4.5 menghitung *entropy* dan *information gain*, dimana loyalitas konsumen adalah atribut tujuan (*class*) dan harga, kemasan, citarasa, variasi, iklan, distribusi, dan kualitas adalah atribut input untuk mendapatkan simpul akar. dan node lainnya [8]. Berdasarkan permasalahan yang sudah dipaparkan di atas maka penelitian ini akan dibuat dengan judul "Prediksi Loyalitas Pelanggan Pada *Fast Moving Consumer Goods* Menggunakan Klasifikasi Metode C4.5" perancangan aplikasi ini diharapkan manager perusahaan dapat mendapatkan data yang efektif, efisien, mudah, dan menguntungkan dalam menentukan strategi.

2. Metode Penelitian

Tahapan dimulai dari studi literatur terlebih dahulu dengan membaca referensi yang terkait dengan penelitian ini selanjutnya merumuskan masalah dari penelitian yang akan dilakukan setelah itu baru dilakukan penyusunan kuisioner kepada pelanggan mie instan untuk menggali informasi terkait penelitian ini. Pada tahap pengumpulan data, dikumpulkan beberapa data yang berhubungan dengan atribut penentu untuk loyalitas pelanggan mie instan, setelah itu tahap analisa sistem dimana pada tahap ini dianalisis kebutuhan sistem yang nanti akan dibuat samapi dengan ke tahap perancangan sistem yaitu dibuat program sesuai dengan rancangan yang telah dibuat. Selanjutnya tahap pengujian hipotesa, dari hasil diatas yaitu tahap implementasi yang sudah dilakukan kemudian dilakukan proses pengujian program untuk menemukan kekurangannya, dan dilakukan proses revisi untuk menemukan hasil yang sesuai dengan persyaratan yang ada. Tahap terakhir adalah kesimpulan dari hasil pengujian hipotesa yang telah dilakukan [9].



Gambar 1. Alur Penelitian

Berkaitan dengan rancangan kegiatan di atas, maka ruang lingkupnya adalah sistem yang akan dibuat menggunakan data mining klasifikasi dengan metode C4.5., data sample yang diolah berjumlah 400 yang berasal dari survey terhadap konsumen FMCG yaitu mie instan dengan 7 atribut yang akan dipakai, diantaranya adalah harga, kemasan, citarasa, variasi, iklan, distribusi, dan kualitas. Sistem yang hendak dibangun berbasis desktop memakai Bahasa pemrograman Java

Data yang peneliti gunakan dalam penelitian ini adalah data kuesioner. Untuk menentukan jumlah data yang dibutuhkan peneliti menggunakan metode Solvin sebagai berikut :

$$n = \frac{N}{1 + N e^2} \quad (1)$$

Keterangan :

n : Jumlah sampel

N : Jumlah Populasi

e = *Margin of Error* (sig. =0.05 atau nilai presisi 95%)

Jumlah penduduk Indonesia berdasarkan Kementrian Perencanaan Pembangunan Nasional (Bappenas) pada tahun 2021 adalah 265 juta jiwa. Maka jumlah sampel dari penelitian ini berdasarkan rumus Slovin tersebut jumlah sampel dari penelitian ini berdasarkan metode Slovin yang digunakan sebesar 399.999 dibulatkan menjadi 400 data.

Data yang didapatkan dari pembagian kuesioner secara online yaitu sebanyak 400 data kuesioner. Data yang dikumpulkan yaitu data kuesioner yang dibagikan kepada pelanggan mie instan di Indonesia dengan atribut nama, jenis kelamin, umur, harga, kemasan, citarasa, variasi, iklan, distribusi, kualitas produk, loyalitas.

Data kuisisioner yang diperoleh dari responden yang berupa pertanyaan tentang kepuasan pelanggan pada perusahaan mie instan, untuk setiap atribut diberikan nilai untuk bisa menentukan puas atau tidak puasnya responden di hitung berdasarkan nilai yang dihasilkan, nilai tersebut sebagai berikut :

1 = STS = Sangat Tidak Setuju

2 = TS = Tidak Setuju

3 = N = Netral

4 = S = Setuju

5 = SS = Sangat Setuju

Nilai dari masing-masing atribut di atas, dikalikan dengan jumlah dimensi yang digunakan. Sehingga menghasilkan nilai Sangat Setuju = 175; Setuju = 140; Netral = 105; Tidak Setuju = 70;

Sangat Tidak Setuju = 35. Jika nilai yang dihasilkan masing-masing atribut 140-175 maka dinyatakan PUAS, Sedangkan nilai yang dihasilkan 1-139, maka dinyatakan TIDAK PUAS.

3. Hasil Dan Pembahasan

Perhitungan Manual

Contoh berikut diberikan dalam kasus yang diuraikan berikut ini untuk membantu menjelaskan algoritme C4.5 pada tabel 1 dibawah ini:

Tabel 1. Contoh Data Sampel

Harga	Kemasan	Citarasa	Variasi	Iklan	Distribusi	Kualitas	Class
SS	S	S	SS	S	SS	S	Ya
S	S	S	N	N	TS	N	Tidak
S	S	S	N	S	SS	S	Ya
S	N	N	S	N	TS	TS	Tidak
SS	N	S	S	S	SS	N	Ya
S	TS	TS	S	S	S	TS	Tidak
SS	N	S	SS	SS	S	S	Ya
N	S	S	SS	S	TS	S	Ya
SS	N	S	S	S	SS	N	Ya
SS	N	S	SS	SS	S	S	Ya
SS	TS	S	SS	S	SS	S	Ya
SS	N	S	S	S	SS	TS	Ya
S	TS	N	S	S	S	TS	Tidak
S	SS	N	SS	SS	S	S	Ya

Tabel 1 adalah contoh data sampel dari data kuisisioner yang nanti akan dipergunakan untuk perhitungan metode C.4.5, dari data tabel 1 diatas selanjutnya yaitu menghitung nilai *entropy* dan *gain*, hasilnya dapat dilihat pada tabel 2 berikut ini :

Tabel 2. Perhitungan *entropy* dan *gain*

Atribut	Jumlah	Ya	Tidak	Entropy	Gain	
Total	14	10	4	0,863120569		
HARGA	STS	0	0	0	0,469565211	
	TS	0	0	0		
	N	1	1	0		
	S	6	2	4		0,918295834
	SS	7	7	0		0
KEMASAN	STS	0	0	0	0,155968102	
	TS	3	1	2		0,918295834
	N	6	5	1		0,650022422
	S	4	3	1		0,811278124
	SS	1	1	0		0
CITRARASA	STS	0	0	0	0,331346037	
	TS	1	0	1		0
	N	3	1	2		0,918295834
	S	10	9	1		0,468995594

VARIASI	SS	0	0	0	0	0,291691997
	STS	0	0	0	0	
	TS	0	0	0	0	
	N	2	1	1	1	
	S	6	3	3	1	
	SS	6	6	0	0	
IKLAN	STS	0	0	0	0	0,371846243
	TS	0	0	0	0	
	N	2	0	2	0	
	S	9	7	2	0,764204507	
	SS	3	3	0	0	
	STS	0	0	0	0	
DISTRIBUSI	TS	3	1	2	0,918295834	0,31957482
	N	0	0	0	0	
	S	5	3	2	0,970950594	
	SS	6	6	0	0	
	STS	0	0	0	0	
	TS	4	1	3	0,811278124	
KUALITAS	N	3	2	1	0,918295834	0,43454914
	S	7	7	0	0	
	SS	0	0	0	0	

Rumus berikut digunakan untuk menghitung total baris kolom *Entropy* pada tabel 2 :

$$Entropy (Total) = \left(-\frac{jml\ ya}{jml\ data} * \log_2 \left(\frac{jml\ ya}{jml\ data} \right) \right) + \left(-\frac{jml\ tidak}{jml\ data} * \log_2 \left(\frac{jml\ tidak}{jml\ data} \right) \right) \quad (2)$$

$$Entropy (Total) = 0.863120569$$

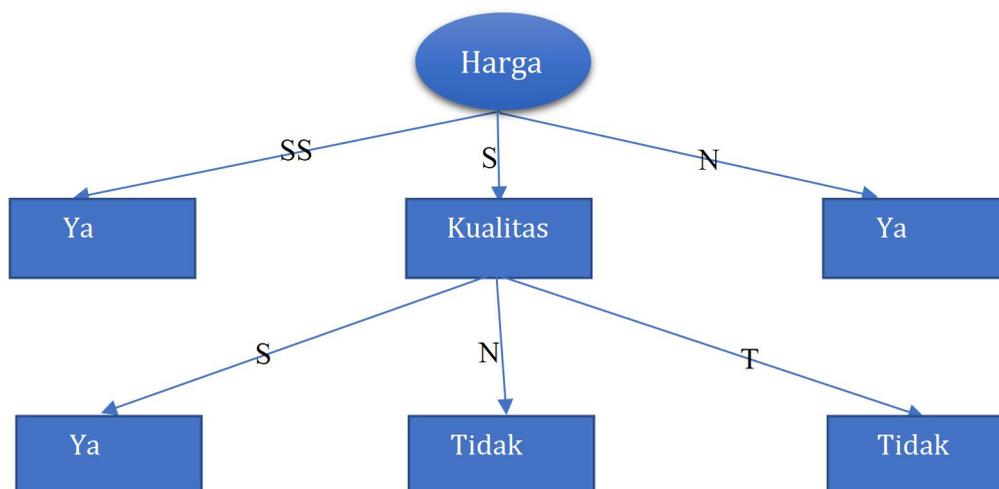
Sedangkan nilai Gain akibat Price dihitung dengan menggunakan rumus Gain, yaitu sebagai berikut:

$$Gain (Total, Harga) = Entropy(Total) - \sum_{i=1}^n \frac{|Harga_i|}{|Total|} * Entropy(Harga) \quad (3)$$

$$Gain (Total, Harga) = 0.4695$$

Dari hasil pada tabel 2 dapat diketahui bahwa atribut dengan Gain tertinggi adalah Harga yaitu sebesar 0.4695. Dengan begitu iklan dapat menjadi node akar. Jadi ada 5 nilai atribut dari Harga yakni STS, TS, N, S, dan SS. Mulai dari 5 nilai atribut yang tertera, nilai atribut N dan SS sudah mengklasifikasikan kasus menjadi 1 yaitu kesimpulan Ya sehingga tidak perlu dilakukan perhitungan lebih lanjut. Atribut N dan SS yang sudah terklasifikasi menjadikan 2 aturan atau rule yang terbentuk, tetapi untuk nilai atribut S masih perlu dilakukan perhitungan lagi karena belum terklasifikasi kasusnya.

Hasil perhitungan sebelumnya dapat ditampilkan dalam bentuk pohon keputusan gambar 2 sebagai berikut:



Gambar 2. Pohon Keputusan Akhir

Dengan hasil pohon keputusan pada gambar diatas, diketahui bahwa semua kasus sudah masuk dalam kelas. Dengan begitu, pohon keputusan pada gambar 2 adalah pohon keputusan terakhir yang terbentuk.

Dari hasil keputusan yang terbentuk hingga terakhir menghasilkan lima aturan (*rules*) atau aturan keputusan dari target yang ingin dicapai yaitu loyal atau tidak loyal. Jadi, hasil *Rules* tersebut diantaranya yaitu :

1. Jika Harga SS maka loyal.
2. Jika Harga N maka loyal.
3. Jika Harga S dan Kualitas S maka loyal.
4. Jika Harga S dan Kualitas N maka tidak loyal.
5. Jika Harga S dan Kualitas TS maka tidak loyal.

Confusion Matrix

Kinerja pendekatan algoritma C4.5 dalam klasifikasi diukur dengan menggunakan *Confusion Matrix* . menggunakan *Confusion Matrix* dapat dihitung secara manual sesuai tabel 3 sebagai berikut:

Tabel 3. Contoh Perhitungan *Confusion Matrix*

	A	B
A = Ya	10	0
B = Tidak	0	4

Berikut ini adalah metode hitungan akurasi dan *error rate* pada data *confusin matrix* berdasarkan tabel:

$$Akurasi = \frac{\text{Bayaknya prediksi Ya}}{\text{Total banyak prediksi}} = \frac{f_{11}+f_{00}}{f_{11}+f_{10}+f_{01}+f_{00}} \quad (4)$$

$$Error Rate = \frac{\text{Bayaknya prediksi Tidak}}{\text{Total banyak prediksi}} = \frac{f_{10} + f_{01}}{f_{11} + f_{10} + f_{01} + f_{00}} \quad (5)$$

$$Akurasi = 1$$

$$\text{Presentase Akurasi} = 1 \times 100\% = 100 \%$$

$$Error Rate = \frac{0+0}{10+0+0+4} = 0$$

Presentase Error = 0 x 100% = 0%

Implementasi

Setelah perancangan yang dilakukan oleh peneliti, maka rancangan tersebut akan dikembangkan menjadi sebuah aplikasi yang berbasis desktop. Aplikasi yang telah dikembangkan mempunyai tampilan sebagai berikut :



Gambar 3. Login



Gambar 4. Beranda

No Data	Nama	Umur	Gender	Harga	Kemasan	Citrarasa	Variasi	Ikan	Distribusi	Kualitas	Class
1	Imam Baih	23	L	SS	S	SS	S	S	SS	S	Ya
2	M Alband	60	L	S	S	S	N	N	TTS	N	Tidak
3	Roudhotul J.	55	P	S	S	S	N	S	SS	S	Ya
4	M Faikhul K.	35	L	S	N	N	S	N	TTS	TTS	Tidak
5	Azzahri Mand.	29	P	SS	N	S	S	S	SS	N	Ya
6	M Mafuhul	33	L	S	TTS	TTS	S	S	S	TTS	Tidak
7	Ongky Inma	24	L	SS	N	S	SS	SS	S	S	Ya
8	Tamjisa Ad.	22	L	N	S	S	SS	S	TTS	S	Ya
9	Left Jayanti	22	P	SS	N	S	S	S	SS	N	Ya
10	Fabrio Hen.	23	L	SS	N	S	SS	SS	S	S	Ya
11	Dita Herlina	20	P	SS	TTS	S	SS	S	SS	S	Ya
12	Erwin Loren	23	L	SS	N	S	S	S	SS	TTS	Ya
13	Khairur Rozza	25	L	S	TTS	N	S	S	S	TTS	Tidak
14	Ego Abdro	25	L	S	SS	N	SS	SS	S	S	Ya
15	Fathul Munir	23	L	S	S	S	N	N	TTS	TTS	Tidak
16	M Bahrul	30	L	S	TTS	TTS	S	S	TTS	S	Tidak
17	Lufiya	40	P	S	S	N	S	N	S	TTS	Ya
18	Aswadi Muz	22	L	N	S	N	S	S	SS	N	Tidak
19	Raza	55	P	M	SS	S	S	S	TTS	S	Ya

Jumlah Data : 400

Gambar 5. Data Training

Mining

All Data : 400 Jumlah Ya : 335 Jumlah Tidak : 65 All Entropy : 0.6402569436810506

Attribut	Jumlah Kelas	Jumlah Ya	Jumlah Tidak	Entropy	Gain	node	gain
Harga TS	8	0	0	0.0		Citrarasa	0.34639832330492706
Harga SS	187	145	42	0.44805200655		Chararasa	0.15447432328769656
Harga S	144	110	34	0.71057695385		Harga	0.29744261770809012
Harga N	21	33	8	0.95144039555	0.21018887987	Isian	0.7624905651102497
Kemasan TS	19	4	15	0.74248759954			
Kemasan SS	11	0	11	0.0			
Kemasan S	3	0	3	0.0			
Kemasan N	184	153	31	0.6542181717			
Kemasan N	193	178	15	0.44727200930	0.08824231423		
Citrarasa TS	25	1	24	0.9717625987			
Citrarasa SS	7	0	7	0.0			
Citrarasa S	11	4	7	0.0			
Citrarasa N	334	321	13	0.2372820935			
Citrarasa N	23	12	11	0.9983596415	0.36828832330		
Varian SS	170	160	10	0.05204608477			
Varian S	192	160	32	0.80161903096			
Varian N	78	66	12	0.81938218467	0.1451658538		
Isian TS	5	3	2	0.91090508445			
Isian SS	63	63	0	0.0			
Isian S	278	270	8	0.6637749392			
Isian N	34	19	15	0.9849271915	0.0261600780		

Gambar 6. Mining

History

Detail Delete

No.	Avg Umur	Tanggal	P	L	Akurasi	Error
1	30	2021-07-06 14:11:02.0	90	77.78	84.21	15.79

Gambar 7. History

Pada halaman *history* menampilkan hasil klasifikasi dengan menggunakan data test sebanyak 19 data kuesioner menghasilkan akurasi 84,21%.

Atribut Berpengaruh

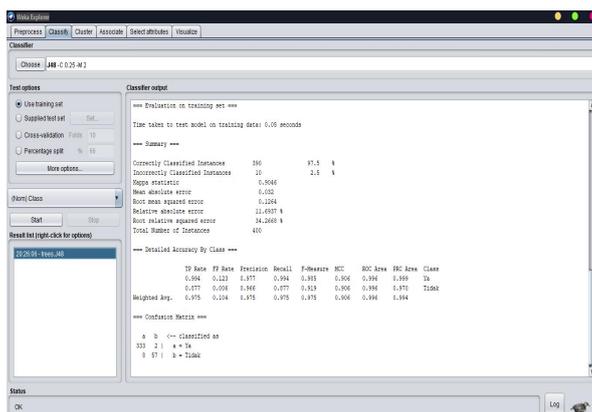
Untuk Meningkatkan loyalitas pelanggan
 Manager membutuhkan sebuah atribut untuk meningkatkan loyalitas pelanggan
 Berikut ini adalah atribut yang telah di temukan

Citrarasa

Gambar 8. Atribut Berpengaruh

Pengujian Metode C.4.5

Data kuesioner yang dikumpulkan kemudian dilakukan pengujian menggunakan Tools WEKA 3.8.5. Sebelum diolah menggunakan WEKA, data dibagi menjadi dua bagian, yang pertama digunakan sebagai data *training* sebanyak 400 data dan yang kedua data *testing* sebanyak 19 data. Selain itu, data yang dibuat dengan menggunakan data *training* akan diuji ulang dengan data testing [10].



Gambar 9. Pengujian Data Training dengan WEKA

Selanjutnya adalah menggunakan WEKA untuk mengevaluasi kinerja hasil klasifikasi Hasil klasifikasi akan diintegrasikan dengan berbagai instrumen pengukuran pada WEKA *classifier* sebagai bukti terjadinya hasil proses klasifikasi yang sudah ada di dalamnya, salah satunya adalah *confusion matrix*. Berikut ini adalah kesimpulan yang dicapai oleh *confusion matrix* berdasarkan data *training*:

```

==== Confusion Matrix ====

a b <-- classified as
333  2 | a = Ya
 8 57 | b = Tidak
    
```

Gambar 10. Confusion Matrix dari Data Training (Sokolova M, 2009).

Proses perhitungan akan dilakukan dengan menggunakan informasi pada gambar di atas, dari persentase akurasi dan *error rate* dalam *confusion matrix* data *training*:

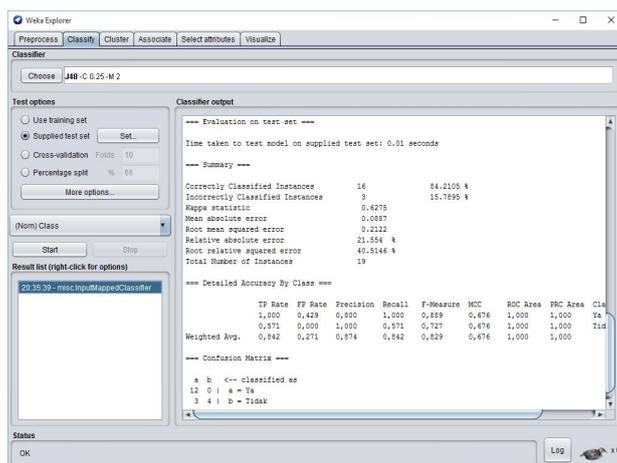
Dengan menggunakan rumus (4) diperoleh akurasi = 0.975

Presentase Akurasi = $0.975 \times 100\% = 97.5\%$

Dengan menggunakan rumus (5) diperoleh error Rate = 0.025

Presentase *error rate* = 2.5%

Langkah selanjutnya adalah mengklasifikasikan data *testing* dengan memanfaatkan Weka untuk mengevaluasi sebanyak 19 data baru untuk pelanggan mie instan. Hasil menggunakan data *testing* ditunjukkan pada grafik di bawah ini:



Gambar 11. Pengujian Data Test dengan WEKA

Hasil yang didapatkan memanfaatkan data perdagangan produk yang dikumpulkan sebagai data *testing* kemudian digunakan untuk menyimpulkan *confusion matrix* yang disediakan dalam aplikasi WEKA.

```

==== Confusion Matrix ====

  a  b  <-- classified as
12  0  | a = Ya
 3  4  | b = Tidak
    
```

Gambar 12. Confusion Matrix dari Data Test

Tabel di bawah ini menunjukkan hasil akurasi dan *error* data *training* dan data *testing*:

Tabel 4. Hasil Pengujian Data menggunakan WEKA

Data set	Akurasi (%)	Error Rate (%)
Data Training	97.5%	2.5%
Data Testing	84%	15%

4. Penutup

Berdasarkan dari penelitian yang dilakukan dengan menggunakan data mining menggunakan algoritma C4.5, dapat dimanfaatkan pada data pelanggan mie instan menjadi kegiatan manajemen strategis berdasarkan penggunaan data mining menggunakan algoritma C4.5 sehingga dapat menampung selama layak dan meningkatkan jumlah pelanggan dengan baik. Hasil dari penelitian menunjukkan bahwa aplikasi dengan menggunakan metode C4.5 menghasilkan akurasi 97.5%, jadi metode C4.5 dapat digunakan untuk membantu manajemen perusahaan mie instan dalam rangka menentukan strategi mempertahankan loyalitas konsumen. Dari 7 atribut yang digunakan yaitu Harga, Kemasan, Citrarasa, Variasai, Iklan, Distribusi, dan Kualitas di dapatkan atribut yang paling berpengaruh adalah atribut

Namun aplikasi ini juga tidak lepas dari kekurangan dan kelemahan. Oleh karena itu, penulis juga memberikan saran untuk pengembangan dan perbaikan sistem selanjutnya, yaitu dengan menambahkan sebuah pohon keputusan (*tree*) secara visual sehingga peneliti dapat mengetahui apa saja atribut yang ada di tingkat *root*, *branch*, dan *leaf*.

5. Referensi

- [1] Indonesia Investment, "Ekonomi Indonesia - Pasar Berkembang Asia | Indonesia Investment," [Online]. Available: <https://www.indonesia-investments.com/id/budaya/ekonomi/item177>.
- [2] BINUS, 2019. [Online]. Available: <https://bbs.binus.ac.id/bbslab/2019/11/%20keadaan-fmcg-di-indonesia/>.
- [3] Tirto.id, Tirto.id, 2016. [Online]. Available: <https://tirto.id/mi-instan-gurih-pasarnya-sengit-persaingannya-yeZ>.
- [4] T. B. Santoso, "Analisa Dan Penerapan Metode C4.5 Untuk Prediksi Loyalitas Pelanggan," *Jurnal Ilmiah Fakultas Teknik LIMIT*, vol. 10, no. 1, 2011.
- [5] P. Chapman, Chapman, P. (2000). CRISP-DM 1.0: Step-by-step data mining guide., 2000.
- [6] F. Gorunescu, *Data Mining Concepts, Models and Techniques*, 2011.
- [7] I. F. Rohman, "Penerapan Algoritma C 4.5 pada Kepuasan Pelanggan Perum DAMRI," *Ilmu Komputer*, pp. 1-14, 2015.
- [8] P. Meilina, "PENERAPAN DATA MINING DENGAN METODE KALSIFIKASI MENGGUNAKAN DECISION TREE DAN REGRESI," *Jurnal Teknologi*, vol. 7, no. 1, 2015.
- [9] J. L. Whitten, L. D. Bentley and K. Dittman, *Metode Desain dan Analisis Sistem*, 2004.
- [10] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification," *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009.
- [11] Indrajani, *Database Design (Case Study All in One)*, PT. Elex Media Komputindo, 2015.
- [12] E. Irfiani and F. Indriyani, "Data Mining untuk Sistem Pengambilan Keputusan Menentukan Kenaikan Kelas Berbasis Web," *Informatics for Educators and Professionals*, vol. 2, no. 1, 2017.
- [13] Kusriani and E. T. Luthfi, *Algoritma Data Mining*, 2009.
- [14] A. Y. R. Pradana, *ANALISIS KEPUASAN PELANGGAN TERHADAP LOYALITAS PELANGGAN MIE SEDAAP (studi kasus pada Indomaret cabang Gedangan Sidoarjo)*, Repository UPN Veteran JATIM, 2011.