

ANALISA ALGORITMA C4.5 UNTUK MEMPREDIKSI PENJUALAN MOTOR PADA PT. CAPELLA DINAMIK NUSANTARA CABANG MUKA KUNING

Nurul Azwanti

Program Studi Sistem Informasi, Fakultas Teknik, Universitas Putera Batam
Jalan R. Soeprapto – Muka Kuning, Batam, Kepulauan Riau, Indonesia.
E-mail : nurulazw@rocketmail.com¹⁾

ABSTRAK

Motor menjadi pilihan setiap orang sebagai andalan transportasi karena dari harga yang terjangkau serta dapat digunakan dalam jangka waktu yang lama. Di Indonesia sendiri, motor yang paling banyak diminati adalah motor merk Honda. Tercatat sebagai pemimpin pasar, dengan kontribusi 72,28 persen dan total penjualan sebesar 2.362.047 unit pada tahun 2016 (menurut Asosiasi Industri Sepeda Motor Indonesia). Untuk wilayah kota Batam perusahaan yang menjual motor Honda salah satunya adalah PT. Capella Dinamik Nusantara Cabang Muka Kuning. Wilayah penjualan sangat mempengaruhi terhadap penjualan motor, selain ke showroom langsung, di beberapa tempat ramai seperti mall bisa menjadi promosi dan penjualan motor karena mudah dijangkau oleh masyarakat. Jumlah data yang sangat banyak akan sangat sulit untuk dianalisa. Analisa diperlukan untuk melihat pola dari data penjualan sehingga dapat menghasilkan prediksi penjualan motor yang nantinya akan berguna untuk pendistribusian motor di beberapa wilayah. Dari data konsumen yang begitu banyak, maka dilakukan Data Mining dengan menggunakan algoritma C4.5. Hasil dari kegiatan mining ini diharapkan dapat memberikan sebuah pohon keputusan untuk melihat pola prediksi perilaku konsumen membeli motor.

Kata Kunci - algoritma c4.5, data mining, klasifikasi, prediksi penjualan motor, pohon keputusan

1. PENDAHULUAN

Strategi pencapaian dalam penjualan membuat persaingan bisnis semakin tajam. Faktor utamanya adalah ketelitian konsumen dalam memilih suatu produk yang tidak hanya dilatarbelakangi oleh harga yang ekonomis, namun dapat membantu kegiatan sehari-hari. Persaingan dunia bisnis ini terjadi pada seluruh perusahaan baik barang maupun jasa, termasuk di perusahaan penjualan motor. Motor menjadi pilihan setiap orang sebagai andalan transportasi karena dari harga yang terjangkau serta dapat digunakan dalam jangka waktu yang lama. Sehingga, tidak mengherankan jika motor yang dinilai praktis dengan harga terjangkau menjadi pilihan masyarakat sebagai alternatif untuk membantu menjalankan berbagai aktivitas dan dampaknya pasar motor nasional pun meningkat [1].

Honda merupakan salah satu perusahaan besar yang bergerak di bidang kendaraan yang sangat maju pesat dan terkenal di masyarakat, khususnya di Indonesia produk yang dikeluarkan Honda sangat ditunggu oleh masyarakat karena kualitas produk dari produk Honda itu sendiri [2]. Menurut data AISI (Asosiasi Industri Sepeda Motor Indonesia) per tahun 2016 Honda masih menjadi pemimpin pasar, dengan kontribusi 72,28 persen dan total penjualan sebesar 2.362.047 unit. Di urutan kedua ada Yamaha dengan sumbangsih 810.317 unit, di mana sebagian jatah 24,08 persen. Kemudian, di posisi ketiga ada Kawasaki yang memperoleh 59.635 unit (1,82 persen). Selanjutnya, di posisi keempat dan kelima ada Suzuki yang berjualan 34.915 unit, sedangkan TVS 1.127 unit.

PT. Capella Dinamik Nusantara Cabang Muka Kuning adalah salah satu perusahaan yang berkecimpung dalam bidang alat transportasi khususnya penjualan sepeda motor merk Honda. Selama ini, PT. Capella Dinamik Nusantara Cabang Muka Kuning cukup dikenal masyarakat kota Batam. Mengingat beragam macam jenis sepeda motor yang ditawarkan, tentunya perusahaan melakukan kegiatan memperkenalkan dan mempopulerkan merk dagang (sepeda motor) yang merupakan suatu syarat untuk berhasilnya perusahaan menjual produk tersebut serta untuk membedakannya dengan produk pesaing. Untuk merebut pangsa pasar dibutuhkan keistimewaan produk yang sesuai dengan selera konsumen dari kualitas, harga, suku cadang serta desain yang baik, dengan demikian produk sepeda motor yang ditawarkan mampu menarik minat konsumen sepeda motor supaya menjadi konsumen yang loyal [3].

Proses pendistribusian motor terkait dengan penjualannya memiliki banyak transaksi. Data transaksi penjualan motor selama satu tahun yaitu pada tahun 2016 menunjukkan sekitar 2.241 unit motor yang terjual baik motor segmen *cub*, *matic* maupun *sport*. Motor segmen *cub* dengan penjualan unit terkecil yaitu 157 unit, kemudian di segmen *matic* yang masih menjadi andalan masyarakat yaitu 1795 unit dan disusul segmen *sport* yaitu 540 unit yang tersebar di beberapa wilayah. Wilayah penjualan sangat mempengaruhi terhadap penjualan motor, selain ke *showroom* langsung, di beberapa tempat ramai seperti *mall* bisa menjadi promosi dan penjualan motor karena mudah dijangkau oleh masyarakat.

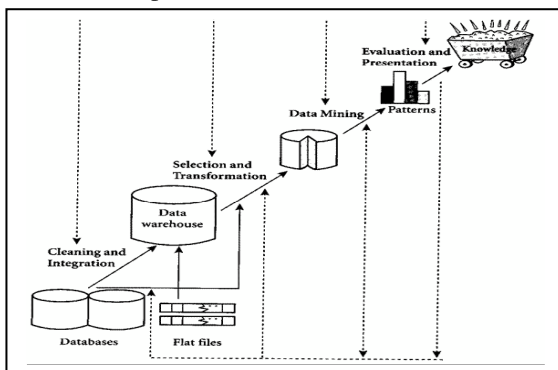
Jumlah data yang sangat banyak akan sangat sulit untuk dianalisa. Analisa diperlukan untuk melihat pola dari data penjualan sehingga dapat menghasilkan prediksi penjualan motor yang nantinya akan berguna untuk pendistribusian motor di beberapa wilayah. Penggalian informasi dari kumpulan data yang berskala besar dapat dilakukan dengan menggunakan teknik *Data Mining*. Salah satu teknik dari *Data Mining* yang dapat digunakan adalah dengan menggunakan algoritma C4.5. Klasifikasi *Data Mining* dengan menggunakan algoritma C4.5 dapat menghasilkan *knowledge* yang dijadikan sebagai dasar pengambilan keputusan dalam memprediksi.

2. TINJAUAN PUSAKA

A. Data Mining

Data mining mengacu pada proses pencarian informasi yang tidak diketahui sebelumnya dari sekumpulan data besar [4]. Pada prosesnya data mining akan mengekstrak informasi yang berharga dengan cara menganalisis adanya pola-pola ataupun hubungan keterkaitan tertentu dari data-data yang berukuran besar [5]. Definisi lain data mining adalah serangkaian proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstrak pengetahuan secara otomatis atau serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual [6].

Karena data mining adalah suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantaraan *knowledge base*. Tahap-tahap ini diilustrasikan pada Gambar 1.



Gambar 1. Tahap-tahap *data mining*

B. Klasifikasi

Klasifikasi data merupakan suatu proses yang menemukan properti-properti yang sama pada sebuah himpunan obyek di dalam sebuah basis data dan mengklasifikasikannya ke dalam kelas-kelas yang berbeda menurut model klasifikasi yang ditetapkan. Tujuan dari klasifikasi adalah untuk menemukan model dari training set yang membedakan atribut ke dalam kategori atau kelas yang sesuai, model tersebut kemudian digunakan untuk mengklasifikasikan atribut yang kelasnya belum diketahui sebelumnya. Beberapa metode klasifikasi umum digunakan dalam

data mining adalah: pengklasifikasi pohon keputusan, pengklasifikasi Bayesian, pengklasifikasi *k-nearest neighbour*, penalaran berbasis kasus, algoritma genetika dan teknik logika *fuzzy* [7].

C. Pohon Keputusan (Decision Tree)

Pohon keputusan menggunakan representasi struktur pohon (tree) di mana setiap node merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut dan daun merepresentasikan kelas. Node yang paling atas dari pohon keputusan disebut sebagai root. Salah satu keuntungan yang paling signifikan dari pohon keputusan adalah kenyataan bahwa pengetahuan dapat diekstrak dan direpresentasikan dalam bentuk klasifikasi aturan if-then [8]. Pohon keputusan merupakan metode klasifikasi yang paling populer digunakan. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami [6]. Pohon keputusan bekerja dengan membentuk pohon keputusan yang dapat disimpulkan aturan-aturan klasifikasi tertentu, salah satunya adalah C4.5 [9].

D. Algoritma C4.5

Algoritma C4.5 yaitu sebuah algoritma yang digunakan untuk membangun decision tree (pengambilan keputusan). Algoritma C.45 adalah salah satu algoritma induksi pohon keputusan yaitu ID3 (Iterative Dichotomiser 3). ID3 dikembangkan oleh J. Ross Quinlan. Dalam prosedur algoritma ID3, input berupa sampel training, label training dan atribut. Algoritma C4.5 merupakan pengembangan dari ID3. Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi missing value, bisa mengatasi kontinu data, dan pruning [10].

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

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

Dimana :

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 entropi dapat dilihat pada persamaan 2 berikut.

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

Dimana :

S : himpunan kasus

A : fitur

N : jumlah partisi S

p_i : proporsi dari S_i terhadap S

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

- Pilih atribut sebagai akar.
- Buat cabang untuk tiap-tiap nilai.
- Bagi kasus dalam cabang.
- Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

3. HASIL PENELITIAN DAN PEMBAHASAN

A. Klasifikasi

Hasil dari data yang telah didapat, diolah dan dilakukan klasifikasi guna untuk melakukan perhitungan.

Tabel 1. Klasifikasi warna

Warna	Klasifikasi
Black Red	BR
Putih Merah Biru	PD
Putih Biru	PB
White	WH
Red Racing	XM
Orange Hitam	OH
Putih Merah	PM
Black Yellow	BY
Putih Pink	PP
White Blue	WL
Black Blue	BB
Biru Silver	LS
Merah Putih	MP
White Red	WR
Hitam Coklat	HF
Krem Coklat	KC
Red Silver	RS
Gray	GR
Red	RD
Black Silver	BS
Putih Hitam	PH
Biru Putih	BP
Merah Coklat	MC
Silver	SV
Merah Silver	MS
Merah Putih	MP
Blue	BL
Black	BG
Putih Orange Merah	PO

Pada klasifikasi selanjutnya yaitu klasifikasi harga, pengelompokan ini diambil harga otr (*on the road*) penjualan motor. Klasifikasi harga dapat dilakukan dengan Rumus Sturgess :

$$\begin{aligned} \text{Jumlah Kelas} &= 1 + 3.3 \log n \text{ (jumlah data)} \\ &= 1 + 3.3 \log (36) \\ &= 6,136 \text{ atau } 6 \end{aligned}$$

Untuk mendapatkan *range* dilakukan perhitungan lagi yaitu:

$$\begin{aligned} \text{Harga Tertinggi} - \text{Harga Terendah} &= 39.754.000 - 13.288.000 = 26.466.000 \text{ kemudian hasilnya dibagi} \\ \text{jumlah kelas} & \end{aligned}$$

$$26.466.000 / 6 = 4.411.000$$

Maka, *range* untuk klasifikasi harga adalah 4.411.000, untuk lebih jelas dapat dilihat pada tabel 2.

Tabel 2. Klasifikasi harga

Harga	Klasifikasi
13.288.000 – 17.699.000	Sangat Terjangkau Sekali (STS)
17.699.001 – 22.110.000	Sangat Terjangkau (ST)
22.110.001 – 26.521.000	Terjangkau (T)
26.521.001 – 30.932.000	Mahal (M)
30.932.001 – 35.343.000	Sangat Mahal (SM)
35.343.001 – 39.754.000	Sangat Mahal Sekali (SMS)

Setelah melakukan klasifikasi, selanjutnya terbentuklah format data akhir yang dapat dilihat pada gambar 2.

No.	SEM.	WARNA	SEGMEN	THN PROD.	HARGA	PENJUALAN	LARIS
1	I	BR	MATIC	2015	STS	SHOWROOM	YA
2	I	PD	SPORT	2015	M	SHOWROOM	TIDAK
3	I	PB	MATIC	2015	STS	SHOWROOM	YA
4	I	WH	MATIC	2015	ST	SHOWROOM	YA
5	I	XM	SPORT	2016	T	SHOWROOM	TIDAK
6	I	OH	CUB	2016	STS	P.PANBL MALL	TIDAK
7	I	PM	MATIC	2016	STS	UMUM	TIDAK
8	I	BY	CUB	2016	STS	SHOWROOM	TIDAK
9	I	PP	SPORT	2015	ST	SHOWROOM	TIDAK
10	II	WL	MATIC	2016	ST	SHOWROOM	YA
11	II	BB	CUB	2016	STS	SHOWROOM	YA
12	II	LS	MATIC	2016	STS	P. SP PLAZA	TIDAK
13	II	MP	SPORT	2016	SM	SHOWROOM	YA
14	II	WR	MATIC	2016	ST	UMUM	TIDAK
15	II	HF	MATIC	2016	STS	P. KEPRI MALL	YA
16	II	WR	MATIC	2016	ST	UMUM	TIDAK
17	II	BY	CUB	2016	STS	SHOWROOM	TIDAK
18	II	KC	MATIC	2016	STS	SHOWROOM	YA
19	III	RS	SPORT	2016	T	P. SP PLAZA	TIDAK
20	III	GR	MATIC	2016	STS	SHOWROOM	YA
21	III	RD	MATIC	2016	STS	SHOWROOM	YA
22	III	BS	SPORT	2016	T	P. KEPRI MALL	YA
23	III	PH	MATIC	2016	STS	P. KEPRI MALL	TIDAK
24	III	LS	MATIC	2016	STS	SHOWROOM	YA
25	III	BP	MATIC	2016	STS	SHOWROOM	YA
26	III	MC	MATIC	2016	STS	SHOWROOM	YA
27	III	SV	MATIC	2016	SMS	SHOWROOM	YA
28	IV	BB	SPORT	2016	ST	P. KEPRI MALL	TIDAK
29	IV	GR	MATIC	2016	STS	P. KEPRI MALL	YA
30	IV	MS	MATIC	2016	STS	SHOWROOM	YA
31	IV	MP	MATIC	2016	STS	SHOWROOM	YA
32	IV	BL	MATIC	2016	ST	SHOWROOM	YA
33	IV	WR	CUB	2016	STS	P. KEPRI MALL	TIDAK
34	IV	BG	MATIC	2016	ST	SHOWROOM	YA
35	IV	PO	SPORT	2016	ST	SHOWROOM	YA
36	IV	PH	MATIC	2016	STS	SHOWROOM	YA

Gambar 2. Format data akhir

Dari *format* data akhir penjualan, maka akan dilakukan klasifikasi data algoritma C4.5 dengan membuat pohon keputusan. Untuk memilih atribut sebagai akar, didasarkan pada nilai *Gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *Gain* digunakan rumus (1), sedangkan untuk menghitung nilai *Entropy* dapat dilihat pada rumus (2), pada tinjauan pustaka sebelumnya.

B. Menghitung Gain dan Entropy

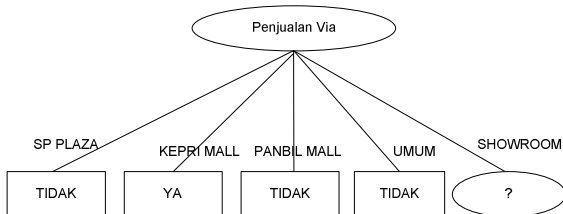
Menentukan atribut sebagai akar dan menghitung nilai informasi Gain atribut. Untuk memilih atribut sebagai akar, didasarkan pada nilai Gain tertinggi dari atribut-atribut yang ada. Dibutuhkan nilai Entropy untuk menentukan Gain tertinggi.

Nilai Entropy dari masing-masing atribut telah didapatkan dan Gain dari atribut juga sudah dihitung, selanjutnya hasil dari perhitungan tersebut dimasukkan ke dalam tabel. Hasil dari perhitungan ini akan dinamakan dengan Node 1.

Node		Jumlah Kasus(S)	Ya (S1)	Tidak (S2)	Entropy	Gain
1	Total	36	21	15	0.9799	
	SEMESTER					
	I	9	3	6	0.9183	0.0820
	II	9	5	4	0.9911	
	III	9	6	3	0.9183	
	IV	9	7	2	0.7642	
	SEGMENT					
	CUB	5	1	4	0.7220	0.1385
	MATIC	23	17	6	0.8281	
	SPORT	8	3	5	0.9544	
	THN PRODUKSI					
	2015	5	3	2	0.9710	0.0001
	2016	31	18	13	0.9812	
	HARGA					
	STS	21	14	7	0.9183	0.1199
	ST	9	5	4	0.9911	
	T	3	1	2	0.9183	
	M	1	0	1	0	
	SM	1	0	1	0	
	SMS	1	1	0	0	
	PENJUALAN VIA					
	P SP PLAZA	2	0	2	0	0.2724
	KEPRIMAL	6	3	3	1	
	P PANBIL	1	0	1	0	
	UMUM	3	0	3	0	
	SHOWROOM	24	18	6	0.8113	

Gambar 3. Perhitungan node 1

Adapun pohon keputusan yang terbentuk dari pencarian Node 1 adalah seperti pada gambar 4.



Gambar 4. Pohon keputusan hasil perhitungan Node 1

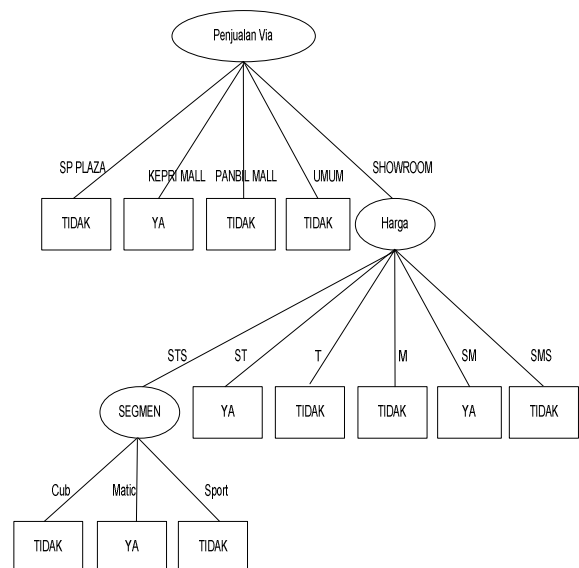
Nilai Entropy dari masing-masing atribut telah didapatkan dan Gain dari atribut juga sudah dihitung, selanjutnya hasil dari perhitungan tersebut dimasukkan ke dalam tabel.

Tabel akhir yang terbentuk dari perhitungan Gain dan Entropy dapat dilihat pada gambar 5.

Node		Jumlah Kasus(S)	Ya(S1)	Tidak (S2)	Entropy	Gain
1.1	Total	14	12	2	0.5917	
	SEMESTER					
	I	3	2	1	0.9183	0.5817
	II	3	2	1	0.9183	
	III	5	5	0	0	
	IV	3	3	0	0	
	SEGMENT					
	CUB	3	1	2	0.9183	0.6965
	MATIC	11	11	0	0	
	SPORT	0	0	0	0	
	THN PRODUKSI					
	2015	2	2	0	0	0.4863
	2016	12	10	2	0.65	

Gambar 5. Perhitungan node 1.2

Dari perhitungan pada gambar 5 dapat diketahui bahwa atribut dengan Gain tertinggi adalah Nilai yaitu **0,6965**. Adapun pohon keputusan akhir yang terbentuk seperti pada gambar 6.



Gambar 6. Pohon Keputusan Hasil Perhitungan Node 1.2

Dari hasil perhitungan node 1.2 tidak ada lagi yang akan menjadi atribut akar, maka perhitungan dihentikan.

Rule hasil dari prediksi berdasarkan pada pohon keputusan terakhir yang terbentuk sesuai dengan perhitungan Entropy dan Gain. Melalui pohon keputusan tersebut diperoleh 12 aturan (rule) dalam memprediksi penjualan motor. Adapun aturan atau rule yang terbentuk dari perhitungan Gain dan Entropy setiap variabel sampai menghasilkan Node 1.2 adalah sebagai berikut :

1. Jika Penjualan Via SP Plaza maka Tidak Laris
2. Jika Penjualan Via Kepri Mall maka Laris
3. Jika Penjualan Via Panbil Mall maka Tidak Laris
4. Jika Penjualan Via umum maka tidak Laris
5. Jika Penjualan Via Showroom dan Harga Sangat Mahal Sekali maka Tidak Laris

6. Jika Penjualan Via Showroom dan Harga Sangat Mahal maka Laris
7. Jika Penjualan Via Showroom dan Harga Mahal maka Tidak Laris
8. Jika Penjualan Via Showroom dan Harga Terjangkau maka Tidak Laris
9. Jika Penjualan Via Showroom dan Harga Sangat Terjangkau maka Laris
10. Jika Penjualan Via Showroom dan Harga Sangat Terjangkau Sekali dan Segmen Cub maka Tidak Laris
11. Jika Penjualan Via Showroom dan Harga Sangat Terjangkau Sekali dan Segmen Matic maka Laris
12. Jika Penjualan Via Showroom dan Harga Sangat Terjangkau Sekali dan Segmen Sport maka Tidak Laris

C. Pengujian

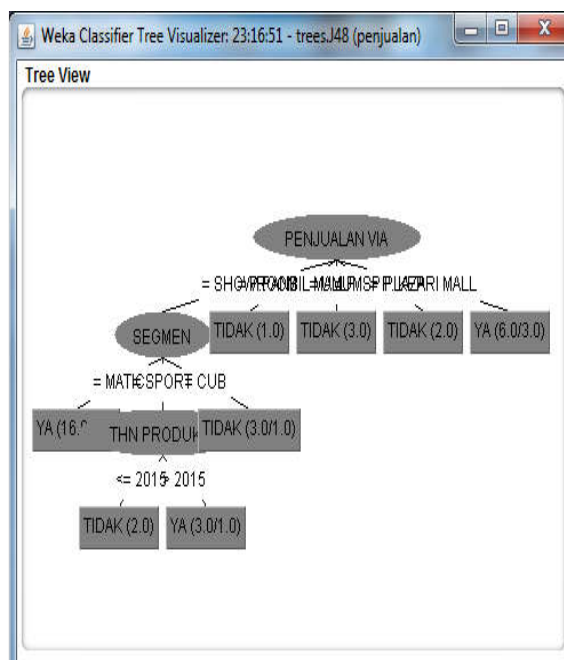
Pengujian terhadap analisa sangat penting dilakukan untuk menentukan dan memastikan apakah hasil analisa tersebut telah sesuai dengan keputusan yang diharapkan. Untuk menguji kebenaran dari hasil pengolahan data yang dilakukan secara manual, maka dapat menggunakan salah satu software aplikasi WEKA 3.8.1.

Adapun langkah-langkah dalam mengolah data menggunakan WEKA adalah sebagai berikut :

Seluruh data variabel dan keputusan yang akan diproses oleh WEKA disimpan terlebih dahulu di *Microsoft Excel* dengan *format .xls*.

Kemudian ubah *format* data tersebut dan *Save as type* cari *format .csv*. Perlu diketahui bahwa *CSV (MS-DOS)* untuk *Window*, *CSV (Macintosh)* untuk pengguna *Apple* dan *CSV (Comma Delimited)* perbedaannya hanya terletak pada koma.

1. Buka *software WEKA 3.8.1*, klik dua kali di *shortcutnya* atau cari melalui *My Computer* dan akan muncul tampilan *WEKA*.
2. Terdapat 4 *button* pilihan *Applications* yaitu *Explorer*, *Experimenter*, *KnowledgeFlow* dan *Simple CLI*. Pilih *Explorer*, kemudian pilih *Open file*, cari dimana *file* penjualan.csv tersebut berada, pilih dan klik *Open*.
3. Langkah selanjutnya pilih variabel yang mempengaruhi data yang akan diproses.
4. Klik menu *Classify*, pada *Classifier* klik *Choose*, untuk algoritma C4.5 pilih *trees* dan klik *J48*.
5. Pilih *Use training set* lalu klik tombol *Start*. Akan tampil *Classifier output* di area sebelah kanan yang merupakan hasil pengolahan dari WEKA. *Classifier output* berisi *running* data yang diinputkan yang merupakan atribut-atribut yang akan membentuk pohon keputusan. Informasi yang ditampilkan berupa jumlah kasus dan keputusannya dan banyaknya cabang dari pohon keputusan.
6. Terakhir, untuk melihat pohon keputusan yang terbentuk dari hasil pengolahan WEKA adalah klik kanan pada *Result list*, pilih *Visualize tree*.



Gambar 7. Visualize tree

Dari *tree view* oleh WEKA, adapun *rule* atau aturan yang terbentuk adalah sebagai berikut :

1. Jika Penjualan via Kepri Mall maka laris
2. Jika Penjualan via SP Plaza maka tidak laris
3. Jika Penjualan via Umum maka tidak laris
4. Jika penjualan via Panbil maka tidak laris
5. Jika penjualan via Showroom dan Segmen Matic maka laris
6. Jika penjualan via Showroom dan Segmen Cub maka tidak laris
7. Jika penjualan via Showroom dan Segmen Sport dan Tahun Produksi 2015 maka tidak laris
8. Jika penjualan via Showroom dan Segmen Sport dan Tahun Produksi 2016 maka laris

Dari dua hasil pengujian yang telah dilakukan yaitu proses secara manual dan menggunakan software WEKA 3.8.1 dapat kita ambil sebuah kesimpulan bahwa hasil pengujian sangat baik karena *rule* yang dihasilkan hampir sama. Perbedaannya hanya terletak pada atribut tahun produksi yang masuk ke dalam WEKA, namun tidak mengubah hasil keputusan.

Algoritma C4.5 dianggap sebagai algoritma yang sangat membantu dalam melakukan klasifikasi data karena karakteristik data yang diklasifikasi dapat diperoleh dengan jelas, baik dalam bentuk struktur pohon keputusan maupun aturan *if-then*, sehingga memudahkan pengguna dalam melakukan penggalian informasi terhadap data yang bersangkutan. Pada pemilihan atribut juga sangat mempengaruhi dalam pengolahan Algoritma C4.5 karena keputusan sangat bergantung pada atribut yang dipilih.

4. KESIMPULAN

Dari hasil penelitian yang telah dilakukan, dapat diambil kesimpulan sebagai berikut :

1. Algoritma C4.5 dapat digunakan sebagai metode klasifikasi dalam memprediksi penjualan motor oleh PT. Capella Dinamik Nusantara Cabang Muka Kuning dengan memperhatikan nilai *gain* (penguatan) tertinggi dari empat atribut seperti Semester, Warna, Tahun Produksi, Segmen dan Harga.
2. Algoritma C4.5 dengan metode pohon keputusan dapat memberikan informasi *rule* prediksi untuk menggambarkan proses yang terkait dengan prediksi penjualan motor.
3. Dari pohon keputusan yang telah terbentuk, variabel tertinggi dalam memprediksi penjualan adalah variabel penjualan via, artinya pendistribusian motor dari berbagai tempat sangat mempengaruhi terhadap penjualan motor. Harga dalam *range* Sangat Terjangkau Sekali (STS) yaitu harga dibawah Rp 17.699.000,00 masih mendominasi motor *matic*. Artinya, *matic* masih menguasai pasar penjualan.
4. Pengujian dengan *WEKA* lebih efektif dan fleksibel karena pohon keputusan yang terbentuk hampir sama dengan perhitungan yang dilakukan dengan manual. Oleh karena itu, penggunaan *WEKA* juga sangat membantu dalam pengujian yang dilakukan.

5. DAFTAR PUSTAKA

- [1] R. Yuliana, "Analisis Strategi Pemasaran Pada Produk Sepeda Motor Matik Berupa Segmentasi, Targeting, Dan Positioning Serta Pengaruhnya Terhadap Keputusan Pembelian Konsumen Di Semarang," vol. 5, no. 2, pp. 79–92, 2013.
- [2] A. Novera, "Pengaruh Biaya Promosi Terhadap Penjualan Sepeda Motor Honda di CV . Mitra Makmur Samarinda," vol. 2, no. 4, pp. 899–913, 2014.
- [3] H. Sastrawinata, "Pengaruh Faktor-Faktor Marketing Stimuli Terhadap Purchasing Decisions Produk-Produk Honda (Studi Kasus Pada PT . Honda Astra Internasional Tbk . Cabang Veteran Palembang," *J. Ilm. Orasi Bisnis*, pp. 122–143, 2010.
- [4] I. H. Selvia Lorena Br Ginting, Wendi Zarman, "Analisis Dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Memprediksi Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik," no. November, 2014.
- [5] N. Iriadi and N. Nuraeni, "Kajian Penerapan Metode Klasifikasi Data Kelayakan Kredit Pada Bank," *J. Tek. Komput. AMIK BSI*, vol. 2, pp. 132–137, 2016.
- [6] A. Sijabat, "Penerapan Data Mining Untuk Pengolahan Data Siswa Dengan Menggunakan Metode Decision Tree (Studi Kasus : Yayasan Perguruan," vol. V, pp. 7–12, 2015.
- [7] H. Chauhan and A. Chauhan, "Implementation of decision tree algorithm c4.5 1," *Int. J. Sci. Res. Publ.*, vol. 3, no. 10, pp. 4–6, 2013.
- [8] G. L. Agrawal and P. H. Gupta, "Optimization of C4 . 5 Decision Tree Algorithm for Data Mining Application," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 3, no. 3, pp. 341–345, 2013.
- [9] D. Untari, "Data Mining Untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree C4.5," 2010.
- [10] S. Faradillah, "Implementasi Data Mining Untuk Pengenalan Karakteristik Transaksi Customer Dengan," pp. 63–70, 2013.