

Pengembangan Aplikasi Perhitungan Prediksi Stock Motor Menggunakan Algoritma C 4.5 Sebagai Bagian dari Sistem Pengambilan Keputusan (Studi Kasus di Saudara Motor)

Pandu Pratama Putra¹, Andi Supriadi Chan²

¹Program Studi Sistem Informasi, Fakultas Komputer STMIK Amik Riau

²Program Studi Sistem Informasi, Fakultas Teknik Dan Komputer Universitas Putera Batam

Email : pandupratamaputra91@gmail.com¹, chakadoang@gmail.com²

Abstrack - The company's resources are limited so that it must be managed effectively and efficiently to ensure the company's viability. Needed an information system technology that will assist management in managing the company so that every decision taken is supported by the actual and reliable information Application of C 4.5 algorithm can be used to predict the motor stock at Saudara Motor. Through the stages of the existing process in the algorithm C 4.5 is preparing the data traning, calculate the roots of the tree, and partition process decision results. The decision process will stop when all the records in node N get the same class, no attribute in the record is in the partition again, there is no record in the empty branch. The result of the calculation of entropy value in algorithm C 4.5 can predict the amount of motor stock provided by the saudara motor dealer. The process of applying Data Mining using C4.5 algorithm obtained decision tree so as to produce 12 rules (rule) in determining prediction of stock stock of motor at dealer of Saudara Motor.

Keywords: dss, algortm c 5.4, motorbike, dealer

Intisari - Sumber daya perusahaan terbatas sehingga harus dikelola dengan efektif dan efisien agar menjamin kelangsungan perusahaan. Dibutuhkan suatu teknologi sistem informasi yang akan membantu manajemen dalam mengelola perusahaan agar setiap keputusan yang diambil ditunjang dengan informasi aktual dan dapat dipercaya Penerapan algoritma C 4.5 bisa digunakan untuk memprediksi stock motor pada Saudara Motor. Melalui tahapan-tahapan proses yang ada pada algoritma C 4.5 yaitu mempersiapkan data traning, menghitung akar dari pohon, dan proses partisi hasil keputusan. Proses hasil keputusan akan berhenti saat semua record dalam simpul N mendapat kelas yang sama, tidak ada atribut di dalam record yang di partisi lagi, tidak ada record di dalam cabang yang kosong. Hasil dari perhitungan nilai entropy pada algoritma C 4.5 dapat memprediksi jumlah stok motor yang disediakan dealer saudara motor. Proses penerapan Data Mining menggunakan algoritma C4.5 diperoleh pohon keputusan sehingga menghasilkan 12 aturan-aturan (rule) dalam menentukan prediksi persediaan stock motor pada dealer Saudara Motor.

Kata Kunci : Sistem Pendukung Keputusan, Algoritma C 4.5, Sepeda Motor, Dealer

I. PENDAHULUAN

Ditengah ketatnya persaingan bisnis yang terjadi belakangan ini sedangkan sumber daya perusahaan terbatas sehingga harus dikelola dengan efektif dan efisien agar menjamin kelangsungan perusahaan maka dibutuhkan suatu teknologi sistem informasi yang akan membantu manajemen dalam mengelola perusahaan agar setiap

keputusan yang diambil ditunjang dengan informasi aktual dan dapat dipercaya. Tujuannya untuk mendapatkan efek seakurat mungkin dari keputusan yang dibuat dan memperkecil ruang kesalahan demi menghemat sumber daya perusahaan.

Pada sistem yang berjalan di Saudara Motor saat sekarang ini masih menggunakan sistem yang manual dalam

pengambilan keputusan pemilihan persediaan stock motor tidak menggunakan data set atau parameter untuk penentuan kriteria unit yang dibutuhkan, sehingga data atau informasi yang tersedia tidak terstruktur dengan baik tapi terstruktur ke arah pohon keputusan. karena pemilihan stock motor yang dilakukan tidak berdasarkan unit apa yang sedang banyak diminati oleh konsumen pada saat sekarang ini. Sehingga dengan cara ini perusahaan mengalami kesulitan dalam pengelompokan terutama pengelompokannya itu harus banyak atau lebih dari sepuluh, jadi aplikasi pada sistem yang lama sulit untuk pengenalan pola dari komponen motor. Tidak jarang stock yang diminta merupakan stock yang peminatnya kurang dan terjadi penumpukan stock di gudang.

Agar dapat menghasilkan keputusan yang tepat untuk pemilihan persediaan stock motor yang sesuai dengan kebutuhan konsumen serta untuk mendukung penjualan pada dealer tersebut, maka penulis akan menggunakan teknik data mining classification dengan metode decision tree induction yang sesuai dengan kebutuhan manajemen operasional dalam menentukan persediaan stock yang memenuhi keinginan konsumen dan dapat mengatasi penumpukan stock. Karena dengan menggunakan metode ini data set yang didapatkan akan di olah sehingga informasi yang tersedia terstruktur kearah pohon keputusan dan mendapatkan pola terbaik untuk penentuan kriteria stock yang diinginkan oleh konsumen.

II. SIGNIFIKASI STUDI

A. Studi Literatur

1. Sistem Pengambilan Keputusan

Decision Support System atau Sistem Pendukung Keputusan merupakan sebuah sistem informasi yang berbasis komputer yang digunakan untuk proses pengambilan keputusan yang menyediakan dukungan informasi yang interaktif bagi menejer dan praktisi bisnis. Dalam sistem ini model yang digunakan adalah model analitis, database, penilaian dan pandangan pembuat

keputusan dan proses pemodelan berbasis komputer yang interaktif guna mendukung pengambilan keputusan yang semi terstruktur[3].

SPK memiliki karakteristik dan kemampuan adalah sebagai berikut :

- a. Mendukung seluruh kegiatan organisasi.
- b. Mendukung beberapa keputusan yang saling berinteraksi.
- c. Dapat digunakan berulang kali dan bersifat konstan.
- d. Terdapat dua komponen utama yaitu data dan model.
- e. Menggunakan baik data eksternal dan internal.
- f. Memiliki kemampuan what-if analysis dan goal seeking Analysis.
- g. Menggunakan beberapa model kuantitatif

2. Knowledge Discovery In Databases

Knowledge Discovery in Database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar[1].

Knowledge Discovery in Database (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses Knowledge Discovery in Database melibatkan hasil proses data mining (proses pengekstrak kecenderungan suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami[8].

Knowledge Discovery in Database (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses Knowledge Discovery in Database melibatkan hasil proses data mining (proses pengekstrak kecenderungan suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami. Ada beberapa macam pendekatan

yang berbeda yang diklasifikasikan sebagai teknik pencarian informasi atau pengetahuan dalam KDD. Ada pendekatan kuantitatif, seperti pendekatan probabilistik seperti logika induktif, pencarian pola, dan analisis pohon keputusan. Pendekatan yang lain meliputi deviasi, analisis kecenderungan, algoritma genetik, jaringan saraf tiruan, dan pendekatan campuran dua atau lebih dari pendekatan yang ada [2].

Pada dasarnya ada enam elemen yang paling esensial dalam teknik pencarian informasi atau pengetahuan dalam KDD yaitu:

- a. Mengerjakan sejumlah besar data.
- b. Diperlukan efisiensi berkaitan dengan volume data.
- c. Mengutamakan ketetapan atau keakuratan.
- d. Membutuhkan pemakaian bahasa tingkat tinggi.
- e. Menggunakan beberapa bentuk dari pembelajaran otomatis.
- f. Menghasilkan hasil yang menarik.

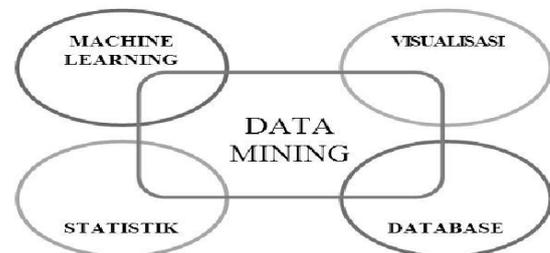
Perusahaan memerlukan kecerdasan bisnis untuk mengembangkan proses bisnis, memonitor waktu, biaya kualitas, dan pengendalian. Gartner Group mendefinisikan kecerdasan bisnis atau business intelligence (BI) sebagai berikut: "Business intelligence is the process of transforming data into information and through discovery transforming that information into knowledge." Dalam definisi ini disampaikan bahwa kecerdasan bisnis merupakan proses perubahan data menjadi informasi. Dari kumpulan informasi yang ada akan diambil polanya menjadi pengetahuan. Tujuan kecerdasan bisnis adalah untuk mengubah data yang sangat banyak menjadi nilai bisnis melalui laporan analistik.

Untuk menggali dan menemukan knowledge diantara setumpukan data yang kelihatannya tidak ada artinya diperlukan teknologi yang telah disebutkan diatas dengan sebutan Data Mining.

3. Data Mining

Menurut Gartner Group data mining adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan, dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika [4].

Data mining menurut Turban adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai database besar. Menurut Gartner Group data mining adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan, dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika [9].



Gambar 1 Data mining Merupakan Irisan dari Berbagai Disiplin

4. Klasifikasi

Sebuah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data dengan tujuan untuk memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Hal ini juga dapat dikatakan sebagai pembelajaran (klasifikasi) yang memetakan sebuah unsur (item) data kedalam salah satu dari beberapa kelas yang sudah didefinisikan [5]

Klasifikasi bertujuan untuk memprediksi target kelas untuk setiap kasus dalam data. Sebuah tugas klasifikasi dimulai dengan satu set data di mana kelas dikenal. Jenis paling sederhana dari masalah klasifikasi adalah klasifikasi biner.

Dalam klasifikasi biner, atribut target hanya memiliki dua nilai yang mungkin. Target multiclass memiliki lebih dari dua nilai [6].

Dalam proses membangun model (pelatihan), algoritma klasifikasi menemukan hubungan antara nilai-nilai prediksi dan nilai target. Algoritma klasifikasi yang berbeda menggunakan teknik yang berbeda untuk menemukan hubungan. Hubungan ini diringkas dalam model, yang kemudian dapat diterapkan pada data yang berbeda ditetapkan dimana kelas tidak diketahui. Klasifikasi model diuji dengan membandingkan nilai-nilai diprediksi nilai target dikenal dalam satu set data uji. Data untuk klasifikasi biasanya dibagi menjadi dua set data: satu untuk membangun model, yang lain untuk pengujian model [10].

Sebuah model klasifikasi diuji dengan menerapkan untuk menguji data dengan nilai target dikenal dan membandingkan nilai prediksi dengan nilai-nilai diketahui. Data uji harus sesuai dengan data yang digunakan untuk membangun model dan harus dipersiapkan dengan cara yang sama. Biasanya data train dan data test berasal dari set data yang sama asalnya. Matrik tes digunakan untuk menilai seberapa akurat model dan memprediksi nilai-nilai yang diketahui. Hasil model klasifikasi berupa kelas dan probabilitas untuk setiap data

5. *Algoritma c.4.5*

Salah satu algoritma yang dapat digunakan untuk membuat pohon keputusan (decision tree) adalah algoritma C4.5. Algoritma C4.5 merupakan algoritma yang sangat populer yang digunakan oleh banyak peneliti di dunia, Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang di ciptakan oleh J. Rose Quinlan [7].

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

- a. pilih atribut sebagai akar
- b. buat cabang untuk tiap-tiap nilai
- c. bagi kasus dalam cabang

- d. ulangi proses untuk setiap 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.

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

Keterangan :

S : Himpunan Kasus A : Atribut

Si : Jumlah Kasus pada Partisi ke-i n :

Jumlah Partisi Atribut A

S| : Jumlah Kasus dalam S

Adapun untuk mencari nilai Entropy, digunakan rumus sbb :

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

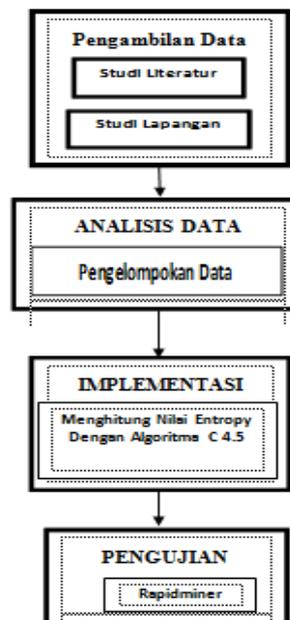
Keterangan

S : Himpunan Kasus A : Fitur

N : Jumlah Partisi S pi : Proporsi dari Si thdp S

B. *Metodologi Penelitian*

Untuk kerangka kerja ini dilakukan beberapa tahapan - tahapan yang harus dilaksanakan mulai dari awal sampai hasil akhir. tahapan kerja tersebut sebagaimana Gambar 1 berikut :



Gambar 1. Kerangka kerja penelitian

III. HASIL DAN PEMBAHASAN

1. Pengambilan Data
Pengambilan data diambil berdasarkan data penjualan yang terjadi selama 1 bulan di tulis dalam bentuk tabel.
2. Analisa Data
Pada tahap ini dilakukan pengelompokan data berdasarkan Type motor dan warna motor di tulis dalam bentuk tabel.
3. Implementasi
Pada tahap ini dilakukan perhitungan pada nilai entropy dengan menggunakan algoritma C 4.5 dari data-data yang di kelompokkan sehingga menghasilkan rule atau pola untuk membangun pohon keputusan
4. Pengujian
Rule yang telah diperoleh dengan Algoritma C 4.5, diujikan dengan Software Data Mining yaitu *Rapidminer*. Proses ini akan menampilkan pola terbaik yang dapat diterapkan untuk menentukan kriteria unit.

A. Analisis data

Data Mining yang diterapkan dalam penelitian ini adalah berbentuk klasifikasi dengan menggunakan algoritma C 4.5 dalam melakukan klasifikasi terhadap data prediksi stock motor. Pengklasifikasian yang dilakukan dengan menggunakan 3 atribut sebagai kriteria dalam menentukan prediksi untuk persediaan stock motor. Atribut yang digunakan dalam penelitian ini adalah :

- a. Type Motor
- b. Jenis Motor
- c. Warna Motor

Output yang dihasilkan dalam penelitian ini terbagi ke dalam 2 class, yaitu prediksi ditambah dan dikurangi. Untuk mencapai tujuan dari penelitian ini, maka dilakukan pengumpulan data. Data sampel yang digunakan adalah berjumlah 35 buah data seperti terlihat pada tabel I.

TABEL I ATRIBUT

JENIS MOTOR	WARNA	TYPE
REVO CW FI	HITAM BIRU	STANDAR DAN MANUAL
REVO FIT FI	HITAM HIJAU	STANDAR DAN MANUAL
VARIO ESP	PUTIH	RINGAN DAN OTOMATIS
VARIO ESP	HITAM	RINGAN DAN OTOMATIS
BEAT SPORTY	PUTIH BIRU	RINGAN DAN OTOMATIS
REVO SPOKE FI	HITAM MERAH	STANDAR DAN MANUAL
BEAT SPORTY	HITAM HIJAU	RINGAN DAN OTOMATIS
REVO FIT FI	HITAM HIJAU	STANDAR DAN MANUAL
CBR150	HITAM	BERAT DAN MANUAL
REVO FIT FI	HITAM HIJAU	STANDAR DAN MANUAL
REVO FIT FI	HITAM HIJAU	STANDAR DAN MANUAL
BEAT POP	HITAM	RINGAN DAN OTOMATIS
REVO CW FI	HITAM	STANDAR DAN MANUAL
BEAT POP	HITAM	RINGAN DAN OTOMATIS

JENIS MOTOR	WARNA	TYPE
REVO CW FI	HITAM MERAH	STANDAR DAN MANUAL
REVO SPOKE FI	HITAM BIRU	STANDAR DAN MANUAL
BEAT POP	HITAM	RINGAN DAN OTOMATIS
REVO FIT FI	HITAM HIJAU	STANDAR DAN MANUAL
CB150R	MERAH	BERAT DAN MANUAL
SCOOPY FI	HITAM MERAH	RINGAN DAN OTOMATIS
REVO CW FI	HITAM BIRU	STANDAR DAN MANUAL
VERZA CW	BIRU	BERAT DAN MANUAL
REVO CW FI	HITAM MERAH	STANDAR DAN MANUAL
REVO SPOKE FI	HITAM HIJAU	STANDAR DAN MANUAL
SCOOPY FI	MERAH HITAM	RINGAN DAN OTOMATIS
VARIO CW	HITAM MERAH	RINGAN DAN OTOMATIS
CB150R	HITAM SILVER	BERAT DAN MANUAL
REVO CW FI	PUTIH MERAH	STANDAR DAN MANUAL
REVO FIT FI	HITAM HIJAU	STANDAR DAN MANUAL
VARIO CW	HITAM MERAH	RINGAN DAN OTOMATIS
VARIO ESP	HITAM	RINGAN DAN OTOMATIS
REVO CW FI	HITAM MERAH	STANDAR DAN MANUAL
REVO SPOKE FI	HITAM BIRU	STANDAR DAN MANUAL
VARIO TECHNO	HITAM SILVER	RINGAN DAN OTOMATIS
VARIO CW	HITAM MERAH	RINGAN DAN OTOMATIS

Setelah data motor dikumpulkan, tahap selanjutnya adalah dilakukan pengolahan data terhadap data yang telah terkumpul tersebut. Proses pengolahan Data Mining yang dilakukan dalam penelitian ini terdiri dari beberapa tahapan diantaranya adalah data selection, data cleaning, data transformation, dan pengolahan Data Mining dengan menggunakan algoritma C 4.5

B. Data Mining Dengan Menggunakan Algoritma c.45

Pada Bagian ini digambarkan arsitektur dari Sistem pendukung Keputusan untuk menentukan karyawan

berprestasi pada KFC batam cabang batu aji

Langkah-langkah pengolahan Data Mining dengan menggunakan algoritma C4.5 adalah sebagai berikut:

1. Pilih atribut sebagai akar.

Data dikelompokkan terlebih dahulu dengan menghitung jumlah kasus, Selanjutnya dihitung nilai Entrophy dan Gain dari semua atribut. Untuk menentukan akar ditentukan dari nilai Gain tertinggi. Hasil pengelompokkan data motor dapat dilihat pada tabel II berikut :

TABEL II PENGELOMPOKKAN DATA MOTOR

Node	Jumlah Kasus	Prediksi	
		Ditambah	Dikurangi
1	Total	35	22
	Type Motor		
	Cub	17	9
	Matic	14	9
	Sport	4	4
	Jenis Motor		
	Beat Series	5	4
	Vario Series	7	4
	Revo Series	12	7
	Supra Series	5	2
	Scoopy Series	2	2
	CB150R	2	2
	CBR150	1	1
	Verza	1	1
	Warna		
	Satu Warna	16	10
	Kombinasi	19	12

2. Buat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Setelah data kita kelompokkan seperti pada tabel 4.6, langkah selanjutnya adalah mencari nilai Entropy total dan juga nilai semua atribut yang digunakan. Kemudian dihitung Gain dari masing-masing atribut yang digunakan.

Berikut perhitungan nilai Entropy dan Gain pada node 1:

a. Nilai Entropy Total

Nilai Entropy total merupakan nilai Entropy yang mewakili dari seluruh jumlah total variabel atribut yang ada. Total kasus pada data sampel dalam penelitian ini adalah berjumlah 35 buah data. Menghitung nilai Entropy total dilakukan berdasarkan nilai total prediksi ditambah (13), dikurangi (22), sedangkan jumlah kasus keseluruhan

adalah sebanyak 35. Perhitungan nilai Entropy total adalah sebagai berikut :

$$Entropy(total) = \left(-\frac{13}{35} * \log_2\left(\frac{13}{35}\right)\right) + \left(-\frac{22}{35} * \log_2\left(\frac{22}{35}\right)\right) = 0.951763$$

b. Nilai Entropy Type Motor

Kriteria Pendidikan terdiri dari 3 nilai yaitu cub, matic dan, sport. Perhitungan nilai Entropy pada atribut pendidikan adalah sebagai berikut :

$$Entropy(Cub) = \left(-\frac{8}{17} * \log_2\left(\frac{8}{17}\right)\right) + \left(-\frac{9}{17} * \log_2\left(\frac{9}{17}\right)\right) = 0.997503$$

$$Entropy(Matic) = \left(-\frac{5}{14} * \log_2\left(\frac{5}{14}\right)\right) + \left(-\frac{9}{14} * \log_2\left(\frac{9}{14}\right)\right) = 0.940286$$

$$Entropy(Sport) = \left(-\frac{4}{4} * \log_2\left(\frac{4}{4}\right)\right) + \left(-\frac{0}{4} * \log_2\left(\frac{0}{4}\right)\right) = 0$$

c. Nilai Entropy Jenis Motor

Atribut yang terdapat pada jenis motor yakni beat series, vario series, revo series, supra series, scoopy series, cb150r, cbr150, verza. Berikut adalah perhitungan nilai Entropy jenis motor :

$$Entropy(Beat) = \left(-\frac{2}{5} * \log_2\left(\frac{2}{5}\right)\right) + \left(-\frac{3}{5} * \log_2\left(\frac{3}{5}\right)\right) = 0.970951$$

$$\text{Entropy(Vario Series)} = \left(-\frac{3}{7} * \log_2 \left(\frac{3}{7}\right)\right) + \left(-\frac{4}{7} * \log_2 \left(\frac{4}{7}\right)\right) = 0.985228$$

$$\text{Entropy(Revo)} = \left(-\frac{5}{12} * \log_2 \left(\frac{5}{12}\right)\right) + \left(-\frac{7}{12} * \log_2 \left(\frac{7}{12}\right)\right) = 0.979869$$

$$\text{Entropy(Supra)} = \left(-\frac{3}{5} * \log_2 \left(\frac{3}{5}\right)\right) + \left(-\frac{2}{5} * \log_2 \left(\frac{2}{5}\right)\right) = 0.970951$$

$$\text{Entropy(Scoopy Series)} = \left(-\frac{0}{2} * \log_2 \left(\frac{0}{2}\right)\right) + \left(-\frac{2}{2} * \log_2 \left(\frac{2}{2}\right)\right) = 0$$

$$\text{Entropy(CB150R)} = \left(-\frac{0}{2} * \log_2 \left(\frac{0}{2}\right)\right) + \left(-\frac{2}{2} * \log_2 \left(\frac{2}{2}\right)\right) = 0$$

$$\text{Entropy(CBR150)} = \left(-\frac{0}{1} * \log_2 \left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} * \log_2 \left(\frac{1}{1}\right)\right) = 0$$

$$\text{Entropy(Verza)} = \left(-\frac{0}{1} * \log_2 \left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} * \log_2 \left(\frac{1}{1}\right)\right) = 0$$

$$\text{Entropy(Kombinasi)} = \left(-\frac{7}{19} * \log_2 \left(\frac{7}{19}\right)\right) + \left(-\frac{12}{19} * \log_2 \left(\frac{12}{19}\right)\right) = 0.949452$$

$$= \text{Entropy(S)} - \sum_{i=1}^n \frac{|Type_i|}{|Total|} * \text{Entropy(Type}_i)$$

$$= 0.951763 - \left(\left(\frac{17}{35} * 0.997503\right) + \left(\frac{14}{35} * 0.940286\right) + \left(\frac{4}{35} * 0\right)\right) = 0.091147$$

$$= \text{Entropy(S)} - \sum_{i=1}^n \frac{|Jenis_i|}{|Total|} * \text{Entropy(Jenis}_i)$$

$$= 0.951763 - \left(\left(\frac{5}{35} * 0.970951\right) + \left(\frac{7}{35} * 0.985228\right) + \left(\frac{5}{35} * 0.970951\right) + \left(\frac{12}{35} * 0.979869\right) + \left(\frac{2}{35} * 0\right) + \left(\frac{2}{35} * 0\right) + \left(\frac{1}{35} * 0\right) + \left(\frac{1}{35} * 0\right)\right) = 0.141348$$

d. Nilai Entropy Warna Motor

Nilai Entropy Golongan terdiri atas 2 nilai yaitu satu warna dan kombinasi. Perhitungan nilai Entropy warna motor adalah sebagai berikut :

$$\text{Entropy(Satu Warna)} = \left(-\frac{6}{16} * \log_2 \left(\frac{6}{16}\right)\right) + \left(-\frac{10}{16} * \log_2 \left(\frac{10}{16}\right)\right) = 0.954434$$

$$= \text{Entropy(S)} - \sum_{i=1}^n \frac{|Warna_i|}{|Total|} * \text{Entropy(Warna}_i)$$

$$= 0.951763 - \left(\left(\frac{16}{35} * 0.954434\right) + \left(\frac{19}{35} * 0.949452\right)\right) = 0.000033$$

Hasil perhitungan nilai Entropy dan Gain node 1 dapat dilihat pada tabel 4.3.

TABEL III TABEL HASIL PERHITUNGAN NODE 1

Atribut	Klasifikasi	Jumlah Kasus		Total	Entropy	Gain
		Ditambah	Dikurangi			
TOTAL		13	22	35	0,9517	
TYPE						
MOTOR	CUB	8	9	17	0,9975	0,091147
	MATIC	5	9	14	0,9402	
	SPORT	4	0	4	0	
JENIS						
MOTOR	Beat Series	2	3	5	0,9709	0,141348
	Vario Series	3	4	7	0,9852	
	Supra Series	3	2	5	0,9709	
	Revo Series	5	7	12	0,9798	
	Scoopy Series	0	2	2	0	
	CB150R	0	2	2	0	
	CBR150	0	1	1	0	
	Verza SATU	0	1	1	0	
WARNA	WARNA	6	10	16	0,9544	0,000033
	KOMBINASI					
	ASI	7	12	19	0,9494	

Dari hasil perhitungan nilai Entropy dan Gain seperti terlihat pada tabel 4.1.dapat diketahui bahwa atribut yang

memiliki Gain tertinggi adalah atribut Jenis Motor yaitu sebesar 0.141348. Dengan demikian, atribut Jenis Motor

adalah atribut yang menjadi node akar. Ada 8 nilai yang terdapat pada atribut Jenis Motor yaitu Beat Series, Vario Series, Scoopy Series, Revo Series, Supra Series, CB150R, CBR150, Dan Verza. Dari semua atribut, nilai pada Scoopy, CB150R, CBR150, dan Verza adalah 0, yaitu prediksi persediaan motor adalah dikurangi, sehingga tidak perlu perhitungan lanjut, tetapi untuk nilai atribut Beat, Vario, Revo, dan Supra masih perlu dilakukan perhitungan lagi.

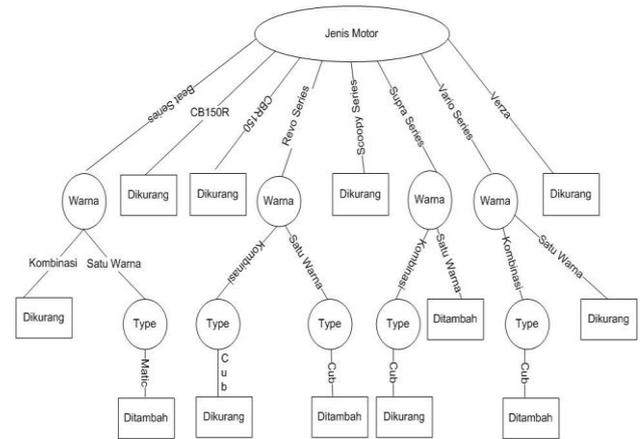
Dari hasil tersebut dapat digambarkan pohon keputusan sementara seperti gambar 4.1.



Gambar 2 Pohon Keputusan Hasil Perhitungan Node 1

Pada gambar 2 di atas dapat diketahui bahwa Jenis Motor cbr150r, cbr150, scoopy, dan verza menyatakan persediaan stock motor dikurangi, sedangkan untuk Jenis Motor beat, revo, supra, vario pada node 1.1, 1.2, 1.3 dan 1.4 harus dilakukan perhitungan lagi.

Dari hasil perhitungan nilai Entropy dan Gain hanya terdapat satu atribut yaitu Type Motor sehingga tidak dapat dilakukan perbandingan nilai Gain. Sehingga pohon keputusan yang terbentuk terlihat pada gambar 3.



Gambar 3 Pohon Keputusan Hasil Perhitungan Node 1

Gambar 3 merupakan hasil akhir dari proses pembuatan pohon keputusan dengan mengimplementasikan Data Mining menggunakan algoritma C4.5 dalam menentukan prediksi stock motor pada dealer Saudara Motor.

Rule-rule yang diperoleh dari penelitian ini adalah sebagai berikut :

1. If Jenis is Beat Series and Warna is Kombinasi Then persediaan is dikurangi.
2. If Jenis is Beat Series and Warna is Satu Warna and Type is Matic Then persediaan is ditambah.
3. If Jenis is CB150R Then persediaan is dikurangi.
4. If Jenis is CBR150 Then persediaan is dikurangi.
5. If Jenis is Revo Series and Warna is Kombinasi and Type is Cub Then persediaan is dikurangi.
6. If Jenis is Revo Series and Warna is Satu Warna and Type is Cub Then persediaan is dikurangi.
7. If Jenis is Scoopy Series Then persediaan is dikurang.
8. If Jenis is Supra Series and Warna is Satu Warna Then persediaan is ditambah.
9. If Jenis is Supra Series and Warna is Kombinasi and Type Is Cub Then persediaan is dikurangi.
10. If Jenis is Vario Series and Warna Is Satu Warna Then persediaan is dikurang.

11. If Jenis is Vario Series and Warna Is Kombinasi and Type Matic Then persediaan is ditambah.
12. If Jenis is Verza Then persediaan is dikurang.

Berdasarkan proses perhitungan yang telah dilakukan melalui beberapa langkah pada proses penerapan Data Mining menggunakan algoritma C4.5 sebelumnya, maka diperoleh pohon keputusan seperti terlihat pada gambar 4.2. Melalui pohon keputusan tersebut diperoleh 12 aturan-aturan (rule) dalam menentukan prediksi persediaan stock motor pada dealer Saudara Motor.

IV. KESIMPULAN

Berdasarkan penelitian dan pembahasan yang dilakukan, maka dapat disimpulkan beberapa hal. Hasil dari perhitungan nilai entropy telah memberikan nilai. Berdasarkan proses perhitungan yang telah dilakukan melalui beberapa langkah pada proses penerapan Data Mining menggunakan algoritma C4.5 maka diperoleh pohon keputusan tersebut diperoleh 12 aturan-aturan (rule) dalam menentukan prediksi jumlah persediaan stock motor pada dealer Saudara Motor.

REFERENSI

- [1] Angga Ginanjar Mabur, L. R. (2012). Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit. *Jurnal Komputer Dan Informatika (KOMPUTA)*, 1(1), 53–57.
- [2] Bidang, P., Sains, K., Mardi, Y., Gajah, J., No, M., & Barat, S. (n.d.). Jurnal Edik Informatika Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . *Jurnal Edik Informatika. Jurnal Edik Informatika*, 213–219.
- [3] Chan, A. S. (2018). Analytical Hierarchy Process Dan Fuzzy Topsis Pada Sistem Pendukung Keputusan, 05(01), 1–14.
- [4] Emha Taufiq Luthfi. (2009). Penerapan Data Mining Algoritma Asosiasi Untuk Meningkatkan Penjualan. *Seminar Nasional Aplikasi Sains & Teknologi (SNAST) Periode III, 10(1)*, 1–21.
- [5] Gunadi, G., & Sensuse, D. I. (2012). Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (FP-GROWTH), 4(1).
- [6] Lorena, S., Zarman, W., & Hamidah, I. (2014). Analisis Dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Memprediksi Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik. *Prosiding Seminar Nasional Aplikasi Sains Dan Teknologi (SNAST)*, (November), 263–272. <https://doi.org/10.5829/idosi.weasj.2015.6.2.22162>
- [7] Nasari, F., Informasi, S., Keputusan, P., & Selection, D. (2014). Penerapan algoritma c4.5 dalam pemilihan bidang peminatan program studi sistem informasi di stmik potensi utama medan, 30–34.
- [8] Tampubolon, K., Saragih, H., Reza, B., Epicentrum, K., Asosiasi, A., & Apriori, A. (2013). Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan. *Informasi Dan Teknologi Ilmiah*, 93–106. <https://doi.org/2339-210X>
- [9] Teknik, F., & Ipa, M. (2015). Kajian Perancangan Rule Kenaikan Jabatan Pada Pt . Abc. *Jurnal SIMETRIS*, 6(2), 217–222.
- [10] Zulkifli, A. (2016). Metode C45 Untuk Mengklarifikasi Pelanggan Perusahaan Telekomunikasi Seluler, 2(1), 65–76.