

Klasifikasi Keuntungan pada Bengkel AS Motor BMW Menggunakan Metode Algoritma C4.5

Julianto ¹, Prima Dina Atika ^{1,*}, Joni Warta ¹

* Korespondensi: e-mail: prima.dina@dsn.ubharajaya.ac.id

¹ Informatika; Universitas Bhayangkara Jakarta Raya; Jl. Raya Perjuangan, Margamulya, Bekasi, Jawa Barat, No. Telp. (021) 88955882; e-mail: julianto18@mhs.ubhara.ac.id, prima.dina@dsn.ubharajaya.ac.id, joniwarta@dsn.ubharajaya.ac.id

Submitted : 13 April 2023

Revised : 26 April 2023

Accepted : 20 Mei 2023

Published : 30 Mei 2023

Abstract

The growth of the motor vehicle and car repair services business is a business that has very high growth in Indonesia, due to the large number of motorized vehicles and cars. Profits in business are important because profits determine whether a company is progressing or not. The method used in this research is Classification Data Mining using the C4.5 Algorithm method, with the attributes involved namely income, cash, spare parts payable, water, electricity, employee salaries, rental costs, consumables, incidental costs, and net income. Algorithm C4.5, which is an algorithm for changing the form of data (tables) into a tree model and then changing the tree model into a rule. The application of the C4.5 algorithm to the profit and loss classification system at workshops was successfully carried out with an accuracy rate of 72.73% from the total data calculation results, and has been implemented in the form of a profit and loss classification application system for website-based workshops. The classification system of advantages and disadvantages in workshops as a result of the research that has been done can be used as a recommendation.

Keywords: Algorithm C4.5, Data Mining, Profit Classification

Abstrak

Pertumbuhan bisnis jasa perbaikan kendaraan bermotor dan bermobil merupakan bisnis memiliki pertumbuhan sangat tinggi di Indonesia, dikarenakan banyaknya kendaraan bermotor dan bermobil. Keuntungan dalam bisnis merupakan hal penting dikarenakan keuntungan menentukan perusahaan mengalami kemajuan atau tidak. Adapun metode yang digunakan dalam penelitian yaitu *Data Mining* Klasifikasi menggunakan metode Algoritma C4.5, dengan atribut yang terlibat yaitu pendapatan, kas, utang sparepart, air, listrik, gaji karyawan, biaya sewa, bahan habis pakai, biaya tak terduga, dan pendapatan bersih. Algoritma C4.5, yaitu algoritma untuk mengubah bentuk data (tabel) menjadi model pohon kemudian mengubah model pohon tersebut menjadi *rule*. Penerapan algoritma C4.5 pada sistem klasifikasi keuntungan dan kerugian pada bengkel berhasil dijalankan dengan menghasilkan tingkat akurasi sebanyak 72,73% dari hasil perhitungan data total, dan telah diimplementasikan dalam bentuk sistem aplikasi klasifikasi keuntungan dan kerugian pada bengkel berbasis website. Sistem klasifikasi keuntungan dan kerugian pada bengkel sebagai hasil dari penelitian yang telah dilakukan dapat dijadikan rekomendasi.

Kata kunci: Algoritma C4.5, *Data Mining*, Klasifikasi Keuntungan

1. Pendahuluan

Banyak orang membicarakan tentang bisnis, keuntungan bisnis serta banyak nya jenis bisnis yang beraneka ragam yang tentu memiliki satu tujuan yaitu mendapatkan keuntungan secara finansial atau laba. Banyaknya usaha dari berbagai sektor yang bersaing untuk mencapai keuntungan yang memuaskan, dan salah satunya usaha dibidang otomotif yang berkembang. Usaha otomotif sekarang sangat menjanjikan dikarenakan setiap masyarakat sudah banyak yang mempunyai kendaraan pribadi beroda empat, yang membutuhkan perawatan *service* atau mengganti *sparepart* yang rusak, sudah pasti membutuhkan jasa bengkel untuk merawat kendaraannya. Masih banyak bengkel ditemukan yang belum terkomputasi dan belum ada pencatatan hasil pendapatan setiap harinya yang menghambat pendataan penghasilan bengkel yang rawan akan kesalahan disebabkan faktor *human eror* menjadi faktor yang menyulitkan dalam pendataan secara manual, susah nya memperkirakan pendapatan dibulan selanjutnya.

Seperti halnya di bengkel AS Motor BMW dalam pembukuan masih menggunakan sistem secara manual, yaitu hanya ditulis dibon atau nota yang sangat rawan untuk pendataan hasil pendapatan. Padahal data pendapatan bengkel sangat bermanfaat dan data yang penting untuk klasifikasi keuntungan laba bengkel untuk meminimalisir kerugian di bengkel AS Motor BMW. Permasalahan lain yang terjadi adalah (a) Data belum terorganisir dengan baik dan masih menggunakan pendataan secara manual dengan menggunakan nota sehingga mempersulit pemilik; (b) Sulitnya memperkirakan atau klasifikasi keuntungan laba dimasa mendatang; dan (c) Belum adanya sistem yang dapat mengklasifikasi keuntungan di bengkel.

Melalui penerapan teknologi *data mining* dan algoritma *C4.5*, sekumpulan data anggaran yang tidak terpakai diolah untuk menghasilkan pengetahuan baru dan menjadi sesuatu yang berharga dan berguna bagi bengkel AS Motor BMW terutama dalam hal peningkatan keuntungan. Tujuan dari penelitian mempelajari bagaimana proses *data mining* menggunakan pendekatan algoritmik *C4.5* membantu untuk memahami pola biaya yang dianggarkan dari bengkel AS Motor BMW saat klasifikasi keuntungan, dan menganalisis serta mengevaluasi aturan yang diturunkan dari pohon keputusan untuk klasifikasi keuntungan.

Penelitian terkait algoritma yang dilakukan oleh peneliti sebelumnya antara lain penerapan algoritma Naïve Bayes pada klasifikasi kelayakan peserta pengajuan kredit rumah di perumahan Azzura Residence diperoleh *accuracy* sebanyak 97.33%, *Precision* Positif/Layak 98.44%, *Precision* Negatif / Tidak Layak sebanyak 90.91%, *Recall* Positif/Layak sebanyak 90.91%, *Recall* Negatif / Tidak Layak sebanyak 98.44% (Karsito & Susanti Santi, 2019). Implementasi algoritma klasifikasi *C4.5* untuk memprediksi kelayakan pembelian kendaraan diperoleh pemilihan kendaraan mobil yang tepat sesuai kebutuhan dengan tingkat keakuratan dari perhitungan 92.4%, *precision* 92.4% dan *recall* 92.4% (Sugiatna et al., 2019). Implementasikan pada prediksi kelahiran bayi secara *premature*, dilihat dari tingkat *accuracy* yang mencapai 97.91 % dan *recall* 98.18%, yang menyatakan bahwa perhitungan yang dilakukan akan mampu memprediksi dan mencegah kelahiran bayi secara *premature* (Jefi, 2019). Algoritma *C4.5* untuk mengetahui tingkat akurasi dalam membuat klasifikasi calon

peserta lomba cerdas cermat siswa SMP. Hasil evaluasi diperoleh bahwa algoritma C4.5 memiliki akurasi 81,81% (Ardiansyah & Walim, 2018). Berdasarkan hasil penelitian-penelitian terdahulu dapat diketahui bahwa implementasi algoritma untuk menyelesaikan permasalahan diperoleh hasil yang akurat.

1.1. Keuntungan

Setiap bisnis bertujuan untuk mencari keuntungan yang maksimal dalam menjalankan kegiatan usahanya, sehingga perusahaan dituntut untuk dapat memanfaatkan sumberdaya yang dimiliki seoptimal mungkin. Persoalan umum yang dihadapi oleh perusahaan adalah bagaimana mengkombinasikan faktor-faktor produksi atau sumberdaya yang dimiliki secara bersama dengan tepat agar diperoleh keuntungan maksimal dengan biaya yang minimal (Indah & Sari, 2019). Secara umum keuntungan dibagi menjadi dua yaitu: (a) Keuntungan dari hasil jual barang, seperti seseorang membutuhkan *sparepart* mobil dari pihak bengkel menyediakan barang yang dibutuhkan pelanggan; dan (b) Keuntungan dari hasil jasa atau keahlian, seseorang membutuhkan perbaikan mobilnya yang rusak dan pelanggan tidak mengerti mesin, maka pihak bengkel menyediakan perbaikan untuk mobil.

1.2. Data Mining

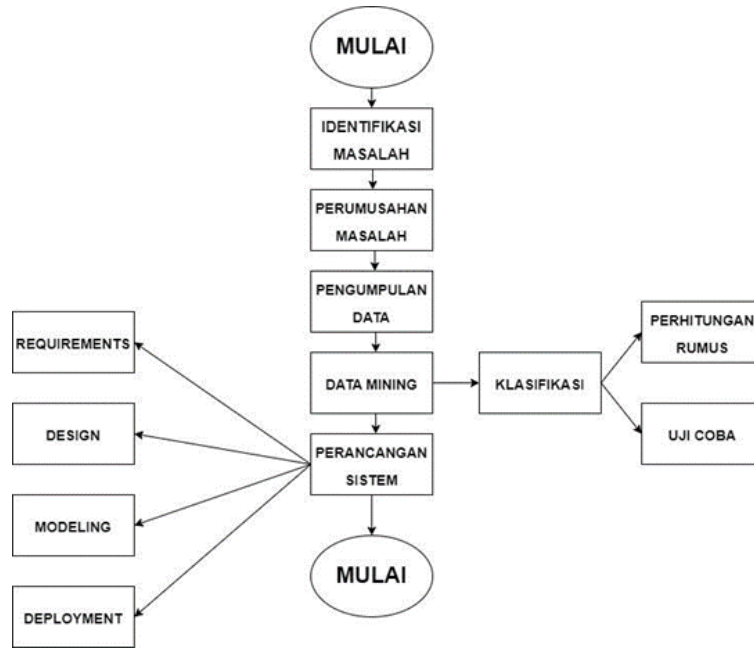
Klasifikasi dalam ilmu data paling menguntungkan, model klasifikasi menunjukkan banyak catatan lengkap yang berisi status *default*, model-model terlatih kemudian disebarkan ke data baru dengan harapan bahwa model yang dibuat akan membantu mengklasifikasikan data (Larose & Larose, 2019).

Data mining didefinisikan sebagai satu set teknik yang digunakan secara otomatis untuk mengeksplorasi secara menyeluruh dan membawa ke permukaan relasi-relasi yang kompleks pada *data set* yang sangat besar. *Data set* yang kompleks pada *data set* data yang berbentuk tabulasi, seperti yang banyak diimplementasikan dalam teknologi manajemen basis data relasional. Akan tetapi, teknik-teknik *data mining* dapat juga diaplikasikan pada representasi data yang lain, seperti domain data spatial, berbasis teks dan multimedia (citra) (Siregar & Puspabhuana, 2017). *Data mining* untuk menemukan pola-pola dalam data. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan.

Karakteristik *data mining* sebagai berikut (a) *Data mining* berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya; (b) *Data mining* biasanya menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk hasil lebih dipercaya; dan (c) *Data mining* digunakan untuk membuat keputusan yang kritis, terutama dalam strategi, juga dapat digunakan untuk pengambilan keputusan dimasa depan berdasarkan informasi yang diperoleh dari data masa lalu.

2. Metode Penelitian

Dalam menyelesaikan penelitian dibutuhkan kerangka pemikiran sebagai pedoman yang dilakukan secara konsisten. Kerangka penelitian ditunjukkan pada Gambar 1. Pendekatan yang dilakukan adalah dengan melakukan menggunakan algoritma C4.5.



Sumber: Hasil Penelitian (2022)

Gambar 1. Kerangka Pemikiran

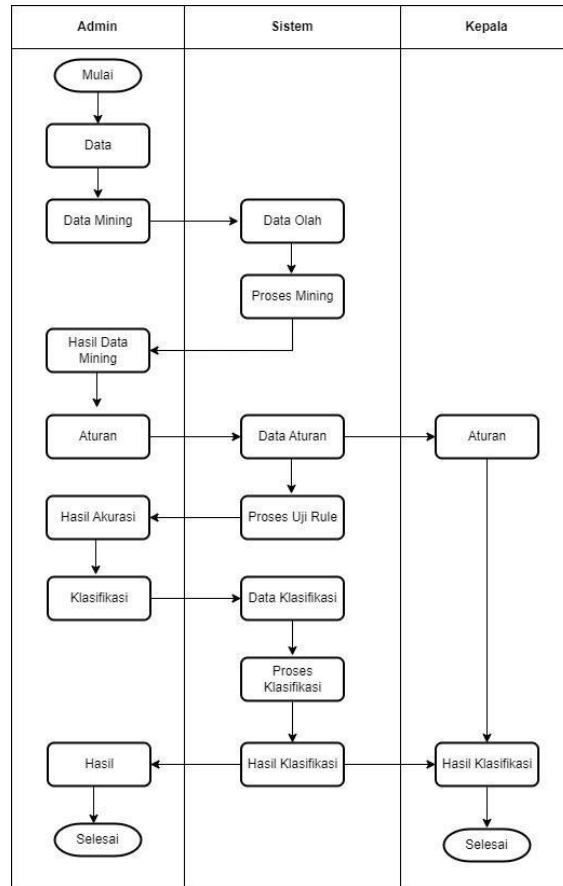
Berdasarkan kerangka pemikiran pada Gambar 1 tahap pertama mengidentifikasi masalah yang terjadi pada bengkel AS Motor BMW. Masalah yang ditemukan dari hasil identifikasi yaitu pendataan secara manual, data menumpuk tidak digunakan untuk menjadi data yang bermanfaat sebagai klasifikasi hasil keuntungan dimasa mendatang. Tahap kedua melakukan perumusan masalah terhadap poin-poin penting yang perlu dilakukan beserta batasannya. Tahap ketiga melakukan pengumpulan data secara primer pengumpulan data secara langsung didapat dari pihak pemilik bengkel. Tahap keempat pengolahan data menggunakan teknik *data mining* dengan mengumpulkan data-data yang ada dari tahun sebelumnya agar bisa diolah untuk perhitungan rumus dan diuji coba menggunakan algoritma C4.5.

2.1. Analisa Sistem Berjalan

Analisis sistem berjalan yaitu adalah suatu proses yang menggambarkan proses yang ada pada objek yang sedang diteliti. Prosedur sistem yang ada pada bengkel AS Motor BMW *customer* melakukan pendaftaran perbaikan mobil kepada admin. Kemudian admin menerima data perbaikan lalu diteruskan ke mekanik untuk dilakukan perbaikan mobil, setelah dilakukan perbaikan admin mencatat perbaikan dan biaya untuk proses pembayaran, laporan diberikan kepada *customer* untuk dibayar. Pimpinan menerima hasil laporan dari admin.

2.2. Analisa Sistem Usulan

Berdasarkan analisis permasalahan dan keadaan sistem berjalan, maka dilakukan pembangunan sistem klasifikasi keuntungan pada bengkel AS Motor BMW berbasis *website* yang dapat membantu pengguna dalam mendapatkan solusi dari permasalahan yang ada, yaitu mempermudah pengolahan data dan untuk mencegah kerugian. Alur sistem usulan pada Gambar 2.



Sumber: Hasil Penelitian (2022)

Gambar 2. Analisis Sistem Usulan

2.3. Algoritma C4.5

Algoritma memiliki karakteristik langkah-langkah yang jelas, pasti dan dapat dijalankan; ditulis dalam Bahasa formal, memiliki nilai awal atau nilai input; menghasilkan keluaran dan proses untuk berhenti (Swastika, 2019). Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau pengelompokan pada *data set*. Dasar dari algoritma C4.5 adalah pembentukan pohon keputusan cabang-cabang dari pohon keputusan merupakan pertanyaan klasifikasi sedangkan untuk daun-daunnya merupakan kelas-kelas atau kelompoknya. Karena tujuan dari algoritma C4.5 untuk melakukan klasifikasi, sehingga hasil dari pengolahan *data set* berupa pengelompokan data ke dalam kelas-kelas tertentu, untuk perhitungan nilai *entropy* menggunakan rumus 1.

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

Keterangan:

- S : himpunan kasus
- A : fitur
- n : jumlah partisi S
- pi : proporsi dari S1 terhadap S

Algoritma dapat membuat pohon keputusan berdasarkan pemilihan atribut yang mempunyai prioritas tertinggi atau juga memiliki nilai *gain* tertinggi berdasarkan nilai *entropy* suatu atribut. Selanjutnya secara berulang-cabang pohon diperluas sehingga seluruh pohon terbentuk. Terdapat empat langkah dalam proses pembuatan pohon keputusan pada algoritma C4.5, yaitu (a) Memilih atribut sebagai akar; (b) Membuat cabang untuk masing-masing nilai; (c) Membagi setiap kasus dalam cabang; dan (d) Mengulangi proses dalam setiap cabang sehingga semua kasus dalam cabang memiliki kelas yang sama untuk selanjutnya dilakukan perhitungan untuk mencari nilai *entropy* dan *gain* (Anggraeni & Andriani, 2021). Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus 2.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Keterangan:

- S : himpunan kasus
- A : atribut
- N : jumlah partisi atribut A
- |S_i| : jumlah kasus pada partisi ke-i
- |S| : jumlah kasus dalam S

2.4. Confussion Matrix

Confusion matrix sebagaimana pada Tabel 1 merupakan pegujian yang dapat digunakan untuk menghitung kinerja atau tingkat kebenaran proses klasifikasi. Dengan *confusion matrix* dapat dianalisa seberapa baik *classifier* dapat mengenali *record* dari kelas kelas yang berbeda (Proboningrum & Sidauruk, 2021).

Tabel 1. Confusion Matrix

Kelas	Terklasifikasi <i>Positive</i>	Terklasifikasi <i>Negatif</i>
<i>Positive</i>	TP (<i>True Positive</i>)	FN (<i>False Negative</i>)
<i>Negative</i>	FP (<i>False Positive</i>)	TN (<i>True Negative</i>)

Sumber: Proboningrum & Sidauruk (2021)

TP adalah *True Positive* yaitu jumlah data *positif* yang terklasifikasi dengan benar oleh sistem, TN adalah *True Negative* yaitu jumlah data *negatif* yang terklasifikasi dengan benar oleh sistem, FN adalah *False Negative* yaitu jumlah data *negatif* namun terklasifikasi salah oleh sistem, FP adalah *False Positive* yaitu jumlah data *positif* namun terklasifikasi salah oleh sistem. Berdasarkan nilai *True Negative (TN)*, *False Negative (FN)*, *True Positive (TP)*, dan *False Positive (FP)* dapat diperoleh nilai akurasi, sensitivitas dan spesifitas.

3. Hasil dan Pembahasan

Penelitian dilakukan di Bengkel AS Motor BMW dari tanggal 1 Maret 2022 sampai 1 Mei 2022 sumber data diperoleh dari dokumen dokumen tertulis yang terdata dipembukuan manual.

3.1. Analisis Data

Analisis data untuk mengumpulkan data yang sudah selesai. Data yang dipakai merupakan data primer. Data primer ialah data yang mempunyai sifat *up to date* dan untuk mendapatkan sebuah data primer peneliti ke lapangan untuk wawancara, observasi dan diskusi dengan pemilik Bengkel AS Motor BMW. *Data selection variabel* yang dipakai pada penelitian data pendapatan bengkel tahun 2021 yang berkaitan dengan pendapatan bersih, dengan pengklasifikasian pendapatan bersih yang dapat dilihat pada Tabel 2.

Tabel 2. Analisis Data Pemasukan dan Pengeluaran

Pemasukan	Pengeluaran	Pendapatan Bersih	
Pendapatan	Kas	Untung	Rugi
	Utang Sparepart		
	Air		
	Listrik		
	Gaji Karyawan		
	Biaya Sewa		
	Bahan Habis Pakai		
	Biaya Tak Terduga		

Sumber: Hasil Penelitian (2022)

Adapun data yang diambil tiga tahun sebelumnya dari 2019-2021 di bengkel dapat dilihat pada Tabel 3.

Tabel 3. Data Bengkel Tahun 2021

Min	Pendap	Kas	Utang	Air	Listrik	Gaji	Biaya	Bahan	Biaya	Penda
ggu	atan		Sparep			Karya	Sewa	Habis	Tak	patan
Ke-			art			wan		Pakai	Terduga	Bersih
	Rp	Rp	Rp	Rp	Rp	Rp	Rp	Rp	Rp	Rp
1	21,000,000.00	1,500,000.00	1,250,000.00	200,000.00	2,400,000.00	4,750,000.00	1,150,000.00	150,000.00	1,600,000.00	8,000,000.00
2	15,000,000.00	1,500,000.00	1,300,000.00	550,000.00	2,300,000.00	4,750,000.00	1,150,000.00	600,000.00	2,100,000.00	750,000.00
3	21,000,000.00	1,500,000.00	900,000.00	450,000.00	1,500,000.00	4,750,000.00	1,150,000.00	250,000.00	1,400,000.00	9,100,000.00
4	11,000,000.00	1,500,000.00	2,000,000.00	200,000.00	2,500,000.00	4,750,000.00	1,150,000.00	100,000.00	1,900,000.00	3,100,000.00
5	12,000,000.00	1,500,000.00	2,400,000.00	350,000.00	1,800,000.00	4,750,000.00	1,150,000.00	700,000.00	2,900,000.00	3,550,000.00
6	20,500,000.00	1,500,000.00	1,800,000.00	150,000.00	1,700,000.00	4,750,000.00	1,150,000.00	550,000.00	1,150,000.00	7,750,000.00
7	11,500,000.00	1,500,000.00	1,750,000.00	600,000.00	1,450,000.00	4,750,000.00	1,150,000.00	400,000.00	2,900,000.00	3,000,000.00

Min ggu Ke-	Pendap atan	Kas	Utang Sparep art	Air	Listrik	Gaji Karya wan	Biaya Sewa	Bahan Habis Pakai	Biaya Tak Terduga	Penda patan Bersih
	Rp	Rp	Rp	Rp	Rp	Rp	Rp	Rp	Rp	Rp
8	19,000,0 00.00	1,500,0 00.00	1,600,0 00.00	400,00 0.00	1,300,0 00.00	4,750,0 00.00	1,150,0 00.00	450,000 .00	1,400,0 00.00	6,450,0 00.00
9	18,000,0 00.00	1,500,0 00.00	1,350,0 00.00	350,00 0.00	1,850,0 00.00	4,750,0 00.00	1,150,0 00.00	100,000 .00	3,800,0 00.00	3,150,0 00.00
10	17,000,0 00.00	1,500,0 00.00	1,450,0 00.00	300,00 0.00	1,750,0 00.00	4,750,0 00.00	1,150,0 00.00	650,000 .00	1,400,0 00.00	4,050,0 00.00
...
...
52	18,800,0 00.00	1,500,0 00.00	1,700,0 00.00	250,00 0.00	1,500,0 00.00	4,750,0 00.00	1,150,0 00.00	400,000 .00	2,950,0 00.00	4,600,0 00.00

Sumber: Hasil Penelitian (2022)

3.2. Analisis Data

Pada proses pengujian dilakukan transformasi data dengan mengklasifikasikan atribut nilai menjadi dua variabel berdasarkan bobot nilai yang didapatkan melalui wawancara kepada pemilik bengkel AS Motor BMW yang dapat dilihat pada Tabel 4.

Tabel 4. Transformasi Data

Bobot Nilai	Kategori
Nilai > 1.500.000	Tinggi
Nilai <= 1.500.000	Rendah

Sumber: Hasil Penelitian (2022)

Atribut pendapatan, kas, utang *sparepart*, air, listrik, gaji karyawan, biaya sewa, bahan habis pakai, dan biaya tak terduga dilakukan pengubahan nilai atribut dari numerik menjadi diskrit (label) pada program sehingga didapatkan label numerik pada Tabel 5.

Tabel 5. Atribut Data Bengkel

Minggu Ke-	Pendapatan	Kas	Utang Sparepart	Air	Listrik	Gaji Karyawan	Biaya Sewa	Bahan Habis Pakai	Biaya Tak Te rduga	Pendapatan Bersih
1	Tinggi	Rendah	Rendah	Rendah	Tinggi	Tinggi	Rendah	Rendah	Tinggi	8,000,000.00
2	Tinggi	Rendah	Rendah	Rendah	Tinggi	Tinggi	Rendah	Rendah	Tinggi	750,000.00
3	Tinggi	Rendah	Rendah	Rendah	Rendah	Tinggi	Rendah	Rendah	Rendah	9,100,000.00
4	Tinggi	Rendah	Tinggi	Rendah	Tinggi	Tinggi	Rendah	Rendah	Tinggi	-3,100,000.00
5	Tinggi	Rendah	Tinggi	Rendah	Tinggi	Tinggi	Rendah	Rendah	Tinggi	-3,550,000.00
6	Tinggi	Rendah	Tinggi	Rendah	Tinggi	Tinggi	Rendah	Rendah	Rendah	7,750,000.00
7	Tinggi	Rendah	Tinggi	Rendah	Rendah	Tinggi	Rendah	Rendah	Tinggi	-3,000,000.00
8	Tinggi	Rendah	Tinggi	Rendah	Rendah	Tinggi	Rendah	Rendah	Rendah	6,450,000.00
9	Tinggi	Rendah	Rendah	Rendah	Tinggi	Tinggi	Rendah	Rendah	Tinggi	3,150,000.00
10	Tinggi	Rendah	Rendah	Rendah	Tinggi	Tinggi	Rendah	Rendah	Rendah	4,050,000.00
...

Minggu Ke-	Pendapatan	Kas	Utang Sparepart	Air	Listrik	Gaji Karyawan	Biaya Sewa	Bahan Habis Pakai	Biaya Tak Terduga	Pendapatan Bersih
...
52	Tinggi	Rendah	Tinggi	Rendah	Rendah	Tinggi	Rendah	Rendah	Tinggi	4,600,000.00

Sumber: Hasil Penelitian (2022)

Selain atribut, kelas atau label *output* yang digunakan juga dilakukan perubahan nilai numerik ke nilai diskrit dengan mencari nilai rata-rata dari total pendapatan bersih pada data tersebut. Setelah didapatkan nilai rata-rata dari data tersebut, ketentuan untuk perubahan nilai numerik ke diskrit adalah: Pendapatan \leq Rp. 1.500.000 : Rugi, sedangkan Pendapatan $>$ 1.500.000 : Laba.

3.3. Proses Pembentukan Pohon Keputusan Menggunakan Algoritma C4.5

Hasil transformasi selanjutnya dianalisis menghasilkan pohon keputusan dengan Algoritma C4.5 formulasi perhitungan *entropy* dan *gain*. Menghitung nilai *entropy* total, nilai *entropy* dari nilai setiap atribut kemudian dihitung nilai *gain* dari setiap atribut. Nilai *entropy* total dan *entropy* dari nilai setiap atribut dihitung menggunakan rumus 1 hasilnya pada Tabel 6. Diketahui jumlah kasusnya adalah 52 kasus dengan kelas terjual tinggi sebanyak 31 dan terjual rendah sebanyak 21.

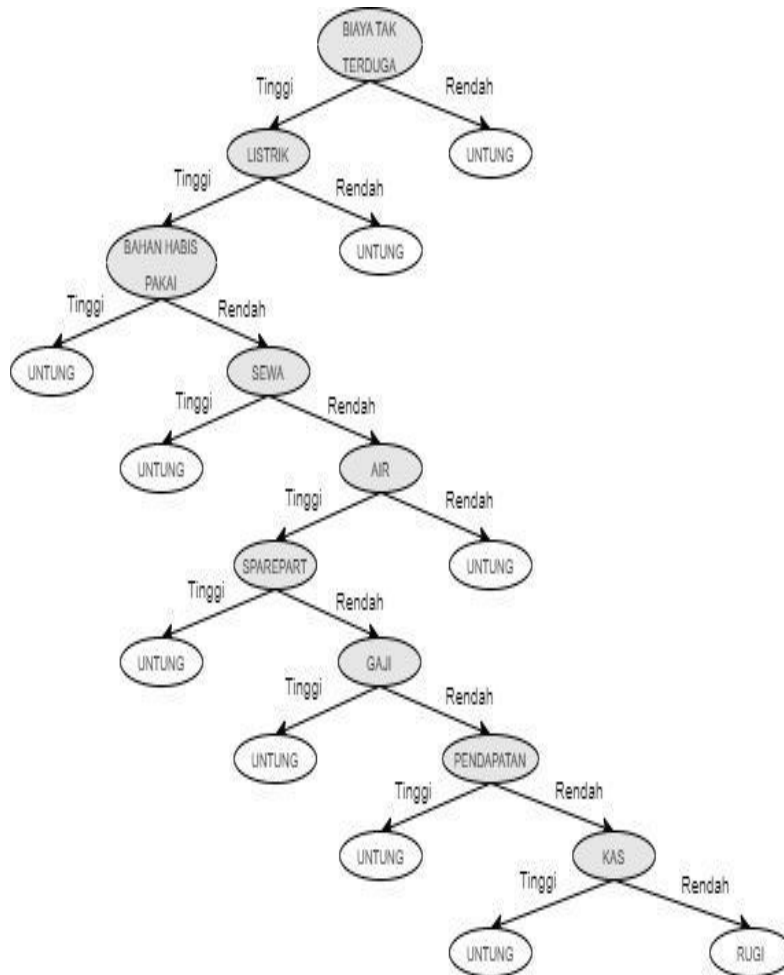
Tabel 6. Hasil Perhitungan Nilai *Entropy* dan *Gain*

Atribut	Class	Jumlah Kasus	Untung	Rugi	Entropy	Gain
Total		52	31	21	0.973	
pendapatan \leq 1500000 pendapatan $>$ 1500000	Rendah	0	0	0	0	
	Tinggi	52	31	21	0.973	0
	Total kasus	52				
kas \leq 1500000 kas $>$ 1500000	Rendah	38	23	15	0.968	
	Tinggi	14	8	6	0.985	0.0007
	Total kasus	52				
sparepart \leq 1500000 sparepart $>$ 1500000	Rendah	21	13	8	0.959	
	Tinggi	31	18	13	0.981	0.001
	Total kasus	52				
air \leq 1500000 air $>$ 1500000	Rendah	36	21	15	0.980	
	Tinggi	16	10	6	0.954	0.001
	Total kasus	52				
listrik \leq 1500000 listrik $>$ 1500000	Rendah	25	18	7	0.855	
	Tinggi	27	13	14	0.999	0.043
	Total kasus	52				
gaji \leq 1500000 gaji $>$ 1500000	Rendah	0	0	0	0	
	Tinggi	52	31	21	0.973	0
	Total kasus	52				
sewa \leq 1500000 sewa $>$ 1500000	Rendah	43	25	18	0.981	
	Tinggi	9	6	3	0.918	0.003
	Total kasus	52				
habis_pakai \leq 1500000 habis_pakai $>$ 1500000	Rendah	37	21	16	0.987	
	Tinggi	15	10	5	0.918	0.006
	Total kasus	52				
tak_terduga \leq 1500000 tak_terduga $>$ 1500000	Rendah	21	16	5	0.792	
	Tinggi	31	15	16	0.999	0.058
	Total kasus	52				

Sumber: Hasil Penelitian (2022)

3.4. Pohon Keputusan

Berdasarkan hasil perhitungan *entropy* dan *information gain* yang didapat. Gambar 3 hasil pohon keputusan berdasarkan hasil perhitungan *gain*. Pada Tabel 6 terlihat atribut biaya tak terduga mempunyai nilai *gain* tertinggi yaitu 0.058 sehingga atribut biaya tak terduga bisa dijadikan sebagai simpul akar (*Root*) dari pohon keputusan.



Sumber: Hasil Penelitian (2022)

Gambar 3. Pohon Keputusan

3.5. Uji Akurasi

Pengujian dilakukan untuk mengetahui kinerja metode algoritma C4.5 dalam melakukan klasifikasi terhadap kelas yang telah ditentukan. Pada pengujian, data yang berjumlah 52 data dibagi menjadi 80% yaitu 41 data *training* dan 20% yaitu 11 data *testing*. Penghitungan *confussion matrix* dari data *testing* pada aplikasi. Jumlah *True Positif (TP)* sebanyak 8 *record*, *False Positif (FP)* sebanyak 0 *record*, *False Negative (FN)* sebanyak 3 *record*, dan jumlah *True Negative (TN)* sebanyak 0 *record*. Akurasi yang dihasilkan setelah melakukan pengujian *confussion matrix*, data *testing* dilakukan uji akurasi, sensitivitas, dan spesifisitas, hasil uji tersebut dapat dilihat pada Tabel 7.

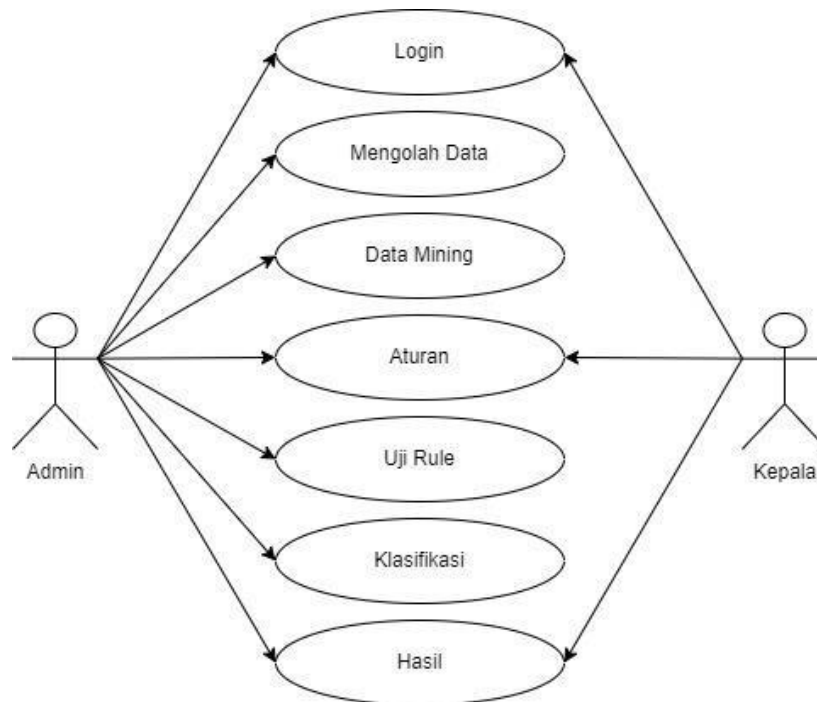
Tabel 7. Hasil Uji Data *Testing*

Uji	Akurasi	Sensitivitas	Spesifisitas
Tingkat Akurasi	72,73%	72,73%	0%

Sumber: Hasil Penelitian (2022)

3.6. Perancangan Sistem dan Implementasi

Menjelaskan tentang ruang lingkup proses sistem klasifikasi keuntungan pada Bengkel AS Motor BMW menggunakan diagram *UML* untuk menggambarkan perancangan sistem, Gambar 4 merupakan diagram *use case* sistem yang dibuat.



Sumber: Hasil Penelitian (2022)

Gambar 4. Diagram *Use Case* Sistem Usulan

Login oleh admin dan kepala bengkel menggambarkan kegiatan pengguna untuk masuk kedalam sistem. Mengolah data oleh admin menggambarkan kegiatan pengguna didalam sistem, yaitu dapat mengupload dan menghapus data. *Data mining* menggambarkan kegiatan Admin didalam sistem yaitu memproses *Data mining*. Pohon keputusan menggambarkan kegiatan admin dan kepala bengkel didalam sistem dapat melihat pohon keputusan, menghapus pohon keputusan. Uji *rule* menggambarkan kegiatan admin didalam sistem melakukan pengujian *rule*. Klasifikasi menggambarkan kegiatan admin didalam sistem dapat menginput data uji. Hasil menggambarkan kegiatan admin dan kepala bengkel didalam sistem melihat *list* hasil pengujian.

Gambar 5 adalah antarmuka sistem klasifikasi keuntungan dengan C4.5, setelah berhasil *login user* akan diarahkan ke Menu Utama (*Dashboard*), yang menampilkan beberapa menu (*Dashboard*, *Data Olah*, *Data Mining*, *Pohon Keputusan*, *Klasifikasi* dan *Hasil*) yang dapat *user* akses.



Sumber: Hasil Penelitian (2022)

Gambar 5. Menu Utama Admin

Halaman Data Olah berfungsi jika admin ingin menambahkan data untuk diupload berupa *file excel* yang bisa dimasukan kedalam program. Setelah memasukan *file excel* maka dimenu *Data mining* akan tampil *file* yang sudah diupload. Kemudian data dicek dan benar dapat dilakukan proses data dengan memilih sub-menu *Proses Mining* seperti pada Gambar 6.

Nilai Atribut	Jumlah data	Jumlah Untung	Jumlah Rugi	Entropy	Gain
pendapatan<=1500000	0	0	0	0	
pendapatan>1500000	41	28	13	0.901	0
kas<=1500000	30	22	8	0.837	
kas>1500000	11	6	5	0.994	0.022
sparepart<=1500000	16	11	5	0.896	
sparepart>1500000	25	17	8	0.904	0
air<=1500000	29	17	12	0.978	
air>1500000	12	11	1	0.414	0.088
listrik<=1500000	19	14	5	0.831	

Sumber: Hasil Penelitian (2022)

Gambar 6. Proses Mining

Dimenu Aturan, admin dapat menguji seluruh hasil untuk mengetahui akurasi dari data yang dimiliki memilih sub-menu Hitung Akurasi untuk menghitung akurasi. Setelah proses menghitung akurasi akan menghasilkan halaman seperti Gambar 7.

Klasifikasi C45											Admin		
6	17200000	1800000	1100000	450000	1350000	4750000	1790000	400000	1900000	untung	untung	11	benar
7	18500000	1500000	1650000	100000	1400000	4750000	1150000	500000	1400000	untung	untung	1	benar
8	11300000	1500000	1800000	2000000	1700000	4750000	1150000	2000000	1200000	rugi	untung	16	salah
9	14700000	1750000	2100000	300000	1500000	4750000	1150000	250000	1650000	untung	untung	11	benar
10	12900000	1500000	1900000	1750000	1900000	4750000	1150000	1500000	1750000	rugi	untung	16	salah
11	10900000	1500000	1250000	1600000	1450000	4750000	1150000	250000	2000000	rugi	untung	13	salah

Jumlah klasifikasi: 11
 Jumlah tepat: 8
 Jumlah tidak tepat: 3

AKURASI = 72.73 %
LAJU ERROR = 27.27 %

 TP=8 TN=0 FP=0 FN=3
 Sensitivitas = 72.73 %
 Spesifisitas = 0 %

Sumber: Hasil Penelitian (2022)

Gambar 7. Hasil Akurasi Uji *Rule*

Dimenu Klasifikasi, admin dapat melakukan Klasifikasi dengan memasukan nilai pada atribut yang ada dan menemukan hasil apakah pendapatan untung atau rugi.

Klasifikasi C45											Admin											
Dashboard											Jumlah data: 3											
Data Olah											No	Pendapatan	Kas	Sparepart	Air	Listrik	Gaji	Sewa	Barang Habis Pakai	Biaya Tak Terduga	Kelas Hasil	Id rule
Data Mining											1	1000000	2000000	2400000	500000	150000	4500000	2400000	1200000	1000000	untung	11
Aturan											2	3000	4000	6000	1000	2000	7000	6000	9000	10000	untung	1
Uji Rule											3	30000	1000	4000	4000	5000	1000	2000	7000	4000	untung	1
Klasifikasi																						
Hasil																						

Sumber: Hasil Penelitian (2022)

Gambar 8. Hasil *Record* dari Klasifikasi

4. Kesimpulan

Aplikasi berbasis Web telah diuji dan dianalisa dalam pembuatan sistem klasifikasi keuntungan *data set* yang dimiliki bengkel AS Motor BMW menggunakan metode algoritma C4.5 dengan menghasilkan tingkat akurasi sebanyak 72,73% dari hasil perhitungan data uji, yang berarti sistem prediksi keuntungan untuk mendeteksi secara benar seluruh data yang diuji sebesar 72,73%. Algoritma C4.5 telah berhasil diuji dan dianalisa dalam pembuatan *rule* untuk sistem klasifikasi keuntungan pada bengkel AS Motor BMW dengan menggunakan data latih berdasarkan variabel: pendapatan, kas, *sparepart*, air, listrik, gaji, sewa, habis pakai, dan tak terduga dengan mendapatkan hasil yang sesuai yang diharapkan, yaitu berdasarkan dua klasifikasi, diantaranya untung dan rugi.

Daftar Pustaka

- Anggraeni, I., & Andriani, S. (2021). Implementasi Algoritma C4.5 untuk Klasifikasi Deteksi Serangan pada Protokol Jaringan. *Komputasi: Jurnal Ilmiah Ilmu Komputer Dan Matematika*, 18(2), 62–68. <https://doi.org/10.33751/komputasi.v18i2.3562>
- Ardiansyah, D., & Walim, W. (2018). Algoritma C4.5 untuk Klasifikasi Calon Peserta Lomba Cerdas Cermat Siswa Smp Dengan Menggunakan Aplikasi Rapid Miner. *Jurnal Inkofar*, 1(2), 5–12.
- Indah, D. R., & Sari, P. (2019). Penerapan Model Linear Programming Untuk Mengoptimalkan Jumlah Produksi Dalam Memperoleh Keuntungan Maksimal (Studi Kasus pada Usaha Angga Perabot). *Jurnal Manajemen Inovasi*, 10(2), 98–115. <https://doi.org/10.24815/jmi.v10i2.16001>
- Jefi. (2019). Prediksi Bayi Lahir Secara Prematur Dengan Menggunakan Metode C4.5 Berbasis Particle Swarm Optimization pada Klinik Umi. *Indonesian Journal on Networking and Security*, 8(2), 1–7.
- Karsito, & Susanti Santi. (2019). Klasifikasi Kelayakan Peserta Pengajuan Kredit Rumah Dengan Algoritma Naive Bayes di Perumahan Azzura Residence. *Jurnal Teknologi Pelita Bangsa*, 9(3), 43–48.
- Larose, C. D., & Larose, D. T. (2019). *Data Science Using Python and R*.
- Proboningrum, S., & Sidauruk, A. (2021). Sistem Pendukung Keputusan Pemilihan Supplier Kain Dengan Metode MOORA. *JSil (Jurnal Sistem Informasi)*, 8(1), 43–48. <https://doi.org/10.30656/jsii.v8i1.3073>
- Siregar, A. M., & Puspabhuana, A. (2017). *DATA MINING: Pengolahan Data Menjadi Informasi dengan RapidMiner*. CV Kekata Group. <https://books.google.co.id/books?id=rTImDwAAQBAJ&printsec=frontcover&hl=id#v=onepage&q&f=false>
- Sugiatna, E., Ibrahim, A. M., & Abdul Hadi, I. (2019). Implementasi Algoritma Klasifikasi C4.5 Untuk Memprediksi Kelayakan Pembelian Kendaraan. *JTIM: Jurnal Teknologi Informasi Dan Multimedia*, 1(2), 124–132. <https://doi.org/10.35746/jtim.v1i2.26>
- Swastika, W. (2019). *Pengantar Algoritma dan Penerapannya pada Python*. Ma Chung Press. https://books.google.co.id/books/about/Pengantar_Algoritma_dan_Penerapannya_Pad.html?id=aSSpDwAAQBAJ&redir_esc=y