

## IMPLEMENTASI ALGORITMA C4.5 UNTUK KLASIFIKASI PRODUK LARIS SEPEDA MOTOR HONDA PADA CV CENDANA MOTOR CEPIRING

<sup>1)</sup> Tabitha Salsabilla, <sup>2)</sup> Sulastri

<sup>1,2)</sup>Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Industri, Universitas Stikubank Semarang

<sup>1,2)</sup>Jl. Tri Lomba Juang, Mugassari – Semarang Selatan – Indonesia

E-mail: [tabithasalsabilla@mhs.unisbank.ac.id](mailto:tabithasalsabilla@mhs.unisbank.ac.id), [sulastri@edu.unisbank.ac.id](mailto:sulastri@edu.unisbank.ac.id)

### ABSTRAK

CV Cendana Motor Cepiring merupakan salah satu perusahaan penjualan sepeda motor merek Honda di Kabupaten Kendal. Persaingan penjualan sepeda motor yang ketat menuntut perusahaan untuk menentukan strategi penjualan yang tepat untuk dapat menaikkan penjualan dan pemasaran produk agar dapat menarik minat para konsumen. Dalam mengetahui ketertarikan konsumen terhadap produk motor Honda, maka dilakukan penelitian mengenai prediksi produk laris sepeda motor Honda dari setiap wilayah kecamatan di Kabupaten Kendal. Metode penelitian yang digunakan adalah algoritma C4.5 *decision tree* dengan prosesnya menggunakan lima langkah pada KDD (*Knowledge Discovery in Databases*). Dari penelitian ini, menghasilkan klasifikasi dengan akurasi sebesar 99% yang menunjukkan bahwa algoritma C4.5 cocok digunakan untuk mengukur perkiraan penjualan sepeda motor Honda terlaris.

**Kata Kunci:** *data mining, klasifikasi, algoritma C4.5, decision tree.*

### ABSTRACT

*CV Cendana Motor Cepiring is one of the Honda brand motorcycle sales companies in Kendal Regency. Tight motorcycle sales competition requires companies to determine the right sales strategy to increase sales and product marketing in order to attract consumers' interest. In knowing consumer interest in Honda motorcycle products, research was conducted on predictions of best-selling Honda motorcycle products from each sub-district in Kendal Regency. The research method used is the C4.5 decision tree algorithm with the process using five steps in KDD (Knowledge Discovery in Databases). This research produces a classification with an accuracy of 99% which indicates that the C4.5 algorithm is suitable for measuring the estimated sales of the best-selling Honda motorcycle.*

**Keywords:** *data mining, classification, C4.5 algorithm decision tree.*

### PENDAHULUAN

Menurut artikel yang diterbitkan oleh Asosisasi Industri Sepeda Motor Indonesia pada 8 November 2021, sepeda motor saat ini menjadi salah satu alat transportasi utama yang diminati di Indonesia. Tercatat angka penjualan selama tahun 2019 hampir menyentuh angka 5 juta unit atau lebih tepatnya 4.910.688 unit, naik dibandingkan tahun sebelumnya yang hanya 4,7 juta unit. Dengan angka tersebut, Honda tercatat sebagai pemimpin pasar dengan kontribusi sekitar 75,69% atau tiga perempat *market share* di

tanah air.

CV Cendana Motor adalah salah satu perusahaan yang bergerak di bidang penjualan sepeda motor merek Honda. Tentunya dengan mempertimbangkan jenis sepeda motor yang ditawarkan, perusahaan berusaha untuk memperkenalkan dan mempopulerkan merek sepeda motor yang merupakan syarat untuk keberhasilan bagi perusahaan yang menjual produk tersebut serta membedakannya dari para pesaing [1]. Tingginya tingkat persaingan dalam dunia bisnis khususnya penjualan motor menuntut *Staff Marketing and Sales* perusahaan untuk mengidentifikasi strategi

yang dapat meningkatkan penjualan dan pemasaran produknya untuk menarik minat para konsumen CV Cendana Motor Cepiring.

Salah satu cara yang dapat dilakukan untuk menentukan strategi adalah dengan menggunakan data penjualan dan data stok persediaan barang untuk diproses dan dianalisis oleh perusahaan untuk membantu perusahaan membuat keputusan bisnis [2].

*Data mining* adalah aktivitas yang mengumpulkan dan menggunakan data historis untuk menemukan pola, keteraturan, dan hubungan dalam kumpulan data yang besar. *Data mining* juga disebut sebagai serangkaian langkah untuk mempelajari nilai tambah secara manual berupa pengetahuan yang tidak diketahui dari suatu kumpulan data [3].

Algoritma C4.5 merupakan kelompok algoritma *decision tree*. Algoritma ini mempunyai input berupa *training samples* dan *samples*. *Training samples* berupa data contoh yang akan digunakan untuk membangun sebuah *tree* yang telah diuji kebenarannya. Sedangkan *samples* merupakan *field – field* data yang nantinya akan digunakan sebagai parameter dalam melakukan klasifikasi data [4].

Pohon keputusan adalah salah satu metode klasifikasi dan prediksi yang kuat dan terkenal dalam aplikasi penambangan data. Pemilihan metode ini didasarkan pada informasi yang ingin diperoleh, khususnya informasi yang bersifat klasifikasi. Keuntungan dari metode ini adalah dapat secara efisien menganalisis sejumlah besar atribut dari data yang ada dan yang bisa dimengerti oleh pengguna akhir [5].

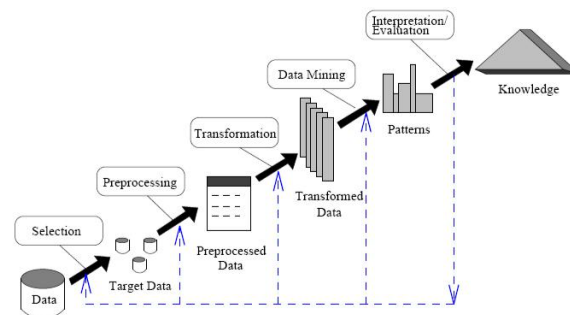
Teknologi pengumpulan dan penyimpanan data telah memudahkan organisasi untuk mengumpulkan data dalam jumlah besar. Ketersediaan data dan informasi yang lengkap, benar dan akurat telah menjadi

kebutuhan mendasar bagi kelangsungan dan berfungsinya suatu organisasi [3]. Melalui *data mining*, perusahaan dapat mengolah data penjualan untuk mengetahui hasil prediksi dengan melihat keakuratan algoritma C4.5 sehingga penjualan sepeda motor Honda dapat memenuhi target sesuai yang telah ditentukan oleh *Staff Marketing and Sales* pada CV Cendana Motor Cepiring.

Menurut Nurul Azwanti (2018), menjelaskan bahwa algoritma C4.5 dapat diimplementasikan untuk melihat model prediksi perilaku konsumsi pengguna saat membeli sepeda motor. Dalam penelitian ini, pihak perusahaan dapat membuat prakiraan penjualan sepeda motor yang akan membantu penjualan dan distribusi di beberapa daerah.

## METODE

Metode penelitian *data mining* yang digunakan untuk menganalisis data dalam penerapan *data mining* ini menggunakan proses tahapan *knowledge discovery in databases* (KDD) [6]:



Gambar 1. Konsep KDD

Tahapan data mining dibagi menjadi 6 (enam) bagian yaitu:

### 1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data penjualan dan persediaan yang diperoleh dari CV Cendana Motor Cepiring. Data ini merupakan data primer yang didapatkan langsung dari perusahaan. Data yang terkumpul sejumlah 6.132 *record* untuk data

penjualan tahun 2020 dan 83 *record* untuk data stok penjualan per Desember 2020.

## 2. Pra Pengolahan Data (*preprocessing*)

Pra pengolahan data adalah proses menghilangkan duplikasi data, mengidentifikasi data yang tidak konsisten, dan mengoreksi kesalahan dalam data [11]. Selain itu, pada proses ini juga dilakukan *enrichment* yaitu proses memperkaya data yang sudah ada dengan data lain yang relevan [7].

## 3. Integrasi Data (*data integration*)

Integrasi data adalah kombinasi data dalam berbagai *database* dalam satu basis data baru. Seringkali data yang diperlukan untuk ekstraksi data disediakan dalam satu *database* serta beberapa *database* atau file teks.

## 4. Seleksi data (*data selection*)

Data dalam basis data yang digunakan seringkali tidak lengkap, sehingga hanya data yang sesuai untuk analisis yang harus diperoleh dari basis data. Misalnya, jika kita ingin melihat kebiasaan belanja orang sebagai bagian dari analisis keranjang belanja, kita tidak perlu menggunakan nama pelanggan, cukup gunakan ID Pelanggan.

## 5. Transformasi data (*data transformation*)

Data diubah atau digabungkan menjadi format yang cocok untuk penambangan data. Beberapa teknik penambangan data memerlukan format data khusus sebelum dapat diterapkan.

## 6. Proses mining

Proses menemukan pola dan informasi yang unik di dalam data yang telah diseleksi sebelumnya dengan menggunakan metode atau teknik tertentu [12].

## 7. Evaluasi pola (*pattern evaluation*)

Proses evaluasi pola bertujuan untuk

menampilkan informasi maupun pola-pola yang telah didapatkan dari proses mining dalam bentuk yang mudah dimengerti oleh *user*. Pada tahap ini didapatkan hasil prediksi dari algoritma C4.5. Selain prediksi juga menghasilkan tingkat akurasi yang didapatkan dari algoritma C4.5 [8].

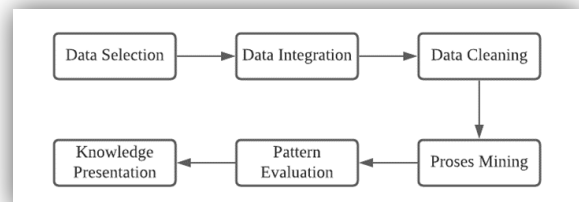
## 8. Presentasi pengetahuan (*knowledge presentation*)

Langkah terakhir dalam proses data mining adalah mengembangkan keputusan atau tindakan berdasarkan hasil analisis. Terkadang orang yang tidak mengerti *data mining* perlu berpartisipasi. Oleh karena itu, menyajikan hasil *data mining* dalam bentuk pengetahuan yang dapat dipahami semua orang merupakan langkah yang diperlukan dalam proses *data mining*. Visualisasi dalam presentasi ini juga dapat membantu mengkomunikasikan hasil *data mining* [2].

## HASIL

### *Penerapan Metode*

Dalam analisis *data mining*, penulis mengikuti Langkah-langkah dalam model *Knowledge Discovery in Databases* (KDD). Dalam penelitian ini hanya memerlukan tahapan sebagai berikut:



Gambar 2. Diagram Alur Penelitian

### *Data Selection*

Pemilihan data adalah langkah memilih atribut data yang akan digunakan untuk pemodelan. Atribut yang digunakan untuk proses *data mining* adalah CusKecamatan, GCust, NamaMotor, MotorWarna, Segmen LeaseKode, dan Keterangan.

### Data Integration

Dalam proses integrasi data, dilakukan penggabungan *database* yaitu data penjualan sepeda motor tahun 2020 dengan data persediaan barang per Desember 2020. Atribut – atribut yang digunakan adalah TypeMotor dan NamaMotor.

### Data Cleaning

Data rekap penjualan dan data persediaan barang yang diperoleh tidak semua *record* berisi data yang lengkap dan juga tidak semua atribut kita butuhkan sehingga perlu dilakukan proses pembersihan data agar diperoleh hasil *data mining* yang berkualitas.

Oleh karena itu, *record* data yang tidak diperlukan seperti *record* kecamatan yang tidak ada di kabupaten Kendal, data kosong atau tidak lengkap dihilangkan dalam proses mining. Sehingga dapat diperoleh *record* data sebanyak 4953 dari data jumlah sebelumnya yaitu 6132 *record* data.

### Classification

Klasifikasi merupakan suatu proses untuk menemukan beberapa model atau fungsi yang menghasilkan gambaran *class* atau konsep dari suatu data [9]. Teknik klasifikasi dengan algoritma C4.5 dalam *data mining* adalah salah satu metode yang dapat dimanfaatkan dalam kaitannya dengan prediksi angka penjualan produk atau barang [10]. Pada tahap klasifikasi ini peneliti menggunakan Rumus *Sturgess* untuk mendapatkan klasifikasi harga dan klasifikasi jumlah terjual [1]:

#### \*Rumus Sturgess

$$\begin{aligned} \text{Jumlah Kelas} &= 1 + 3,3 \log n & [1] \\ &= 1 + 3,3 \log (5.055) \\ &= 1 + 3,3 \log (3,703) \\ &= 1 + 1,875 = 2,875 \text{ atau } 3 \end{aligned}$$

Untuk mendapatkan kisaran klasifikasi harga, dilakukan perhitungan lagi:

#### \*Range Harga

Harga Tertinggi – Harga Terendah =  $40.370.000 - 15.140.000 = 25.230.000$ , kemudian hasilnya dibagi dengan jumlah kelas,  $25.230.000 / 3 = 8.410.000$ , maka *range* untuk klasifikasi harga yaitu 8.410.000. Rincian lebih lanjut dapat dilihat di klasifikasi Harga pada tabel 1.

Tabel 1. Klasifikasi Harga

Harga	Klasifikasi
15.140.000 – 23.550.000	Terjangkau (T)
23.550.000 – 31.960.000	Mahal (M)
31.960.000 – 40.370.000	Sangat Mahal (SM)

Perhitungan klasifikasi jumlah terjual juga sama menggunakan jumlah kelas yaitu 3 dan perhitungan range jumlah terjual:

#### \*Range Jumlah Terjual

Jumlah Jual Banyak – Jumlah Jual Sedikit =  $836 - 1 = 835$ , kemudian hasilnya dibagi dengan jumlah kelasnya,  $835 / 3 = 278,3$ , maka range untuk klasifikasi jumlah terjual yaitu 278. Rincian lebih lanjut dapat dilihat di klasifikasi jumlah terjual pada tabel 2.

Tabel 2. Klasifikasi Jumlah Terjual

Jumlah Terjual	Klasifikasi
1 - 279	Sedikit (S)
279 - 557	Banyak (B)
557 - 836	Sangat Banyak (SB)

Setelah mendapatkan klasifikasi jumlah terjual maka akan dilakukan klasifikasi untuk mendapatkan variabel target yaitu klasifikasi keterangan dengan membandingkan dengan tabel klasifikasi jumlah terjual. Lebih jelasnya bisa dilihat pada tabel 3 klasifikasi keterangan.

Tabel 3. Klasifikasi Keterangan

Klasifikasi Jumlah Terjual	Klasifikasi Keterangan
S	Tidak Laris (TL)
B dan SB	Laris (L)

Tabel 4. Data Setelah di Klasifikasi

No	CusKecamatan	NamaMotor	MotorWarna	DKHarga	JmlTerjual	LeaseKode	Keterangan
1	Gemuh	ANSCOOPY	Black	T	SB	ADIRA	LARIS
2	Cepiring	BSPOCBS P	Merah Hitam	T	S	TUNAI	TL
3	Gemuh	BSPOCBS P	Merah Putih	T	S	FIF	TL
4	Gemuh	BSPOCBS P	Merah Hitam	T	S	ADIRA	TL
...	...	...	...	...	...	...	...
4952	Weleri	V150 SPRT	Hitam Merah	M	S	FIF	TL
4953	Weleri	V150 3S	Silver	M	S	TUNAI	TL

### Data Transformation

Dalam proses transformasi terdapat beberapa atribut yang dapat disederhanakan menjadi sebuah atribut baru. Atribut DKHarga, JumlahTerjual, dan Keterangan

ditransformasikan menjadi atribut Keterangan yang mempunyai *record value* yaitu Laris dan Tidak Laris. Adapun hasil transformasi data diatas ditunjukkan pada gambar berikut ini:

Tabel 5. Transformasi Data

No	CusKec	GCust	NamaMotor	MotorWarna	Segmen	LeaseKode	Keterangan
0	Weleri	P	ADV 150 A	Putih Hitam	Matic	Tunai	Tidak Laris
1	Patebon	P	ADV 150 A	Merah Hitam	Matic	Tunai	Tidak Laris
2	Ngampel	L	ADV 150 A	Putih Hitan	Matic	Tunai	Tidak Laris
3	Kendal	L	ADV 150 A	Merah Hitam	Matic	Tunai	Tidak Laris
...	...	...	...	...	...	...	...
4952	Cepiring	L	V 150 SPRT	Hitam Merah	Matic	FIF	Tidak Laris
4953	Boja	L	V 150 SPRT	Hitam Merah	Matic	FIF	Tidak Laris

### Implementasi Decision Tree

Pengujian terhadap analisa sangat penting dilakukan untuk menentukan dan memastikan apakah hasil analisa tersebut sesuai dengan keputusan yang diharapkan [1]. Pada tahapan ini, proses data mining menerapkan algoritma C4.5 *decision tree* dengan menggunakan Bahasa pemrograman Python dengan *tools* Google Colab dengan beberapa tahapan sebagai berikut:

Data yang sudah siap di-mining dan akan diproses di Google Colab disimpan terlebih dahulu dengan format .csv. perlu

diketahui yang digunakan adalah CSV (Comma Delimited).

- Buka <https://colab.research.google.com/> website. Pada website akan ada button File. Pada File akan ada button 'Bagian Baru' klik button tersebut lalu akan muncul lembar kerja baru. Setelah muncul, klik tanda +Kode untuk mulai menulis perintah.
- Pada File terdapat button 'Upload ke penyimpanan sesi' klik button tersebut untuk mengimport file dataolah8.1.csv.
- Setelah itu ketikkan *script library* yang

dibutuhkan untuk membuat model algoritma C4.5. Disini penulis memakai *library* numpy, pandas, seaborn, dan matplotlib.pyplot.

```
[1] #Import library numpy, pandas, seaborn dan matplotlib.pyplot
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

Gambar 3. *Import Library*

d. Langkah selanjutnya memasukkan *dataset* yang akan digunakan untuk membuat

```
#Memasukkan library untuk membuat model algoritma C4.5 dan akurasiya
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
from sklearn.tree import DecisionTreeClassifier
```

Gambar 4. *Library Model Algoritma C4.5*

h. Masukkan *code* untuk *split* data ke *training set* dan *testing set*. Menerapkan metode ini dapat memberikan hasil prediksi yang lebih akurat untuk data baru atau yang tidak

```
#Kode untuk split data train dan data test
training_data, testing_data = train_test_split(dataset, test_size=0.3, random_state=25)

print(f"No. of training examples: {training_data.shape[0]}")
print(f"No. of testing examples: {testing_data.shape[0]}")

No. of training examples: 3467
No. of testing examples: 1487
```

Gambar 5. Hasil dari *Split Data*

i. Setelah melakukan *split* data, selanjutnya akan menggunakan *code* program untuk menghitung tingkat akurasi algoritma C4.5.

*decision tree*. *Dataset* disini yaitu *file* 'dataolah8.1.csv'.

e. Selanjutnya ubah data *type* dari data *type string* menjadi *type numeric* agar data bisa dibaca oleh Bahasa pemrograman Python.

f. Setelah diubah menjadi *type numeric*, maka variabel 'Keterangan akan di *drop out* untuk dijadikan sebagai variabel *class*.

g. Selanjutnya masukkan beberapa *library* untuk membuat algoritma C4.5 dan akurasiya.

pernah dilatih. Pada tahap ini menggunakan perbandingan 70:30. 70% data train dengan 3.467 data dan 30% data test dengan 1.487 data.

j. Terakhir, yaitu *code* untuk membuat grafik *decision tree* dengan memasukkan *library* 'sklearn.tree'.

```
#Membuat grafik decision tree
#Memasukkan library untuk membuat grafik
from sklearn.tree import export_graphviz
export_graphviz(tree_dataset, out_file="Tree_keterangan.dot",
                class_names=["Laris", "Tidak Laris"],
                feature_names=atr_dataset.columns, impurity=False, filled=True)

#Kode program untuk mengimport grafik
import graphviz

with open("Tree_keterangan.dot") as fig:
    dot_graph = fig.read()
    graph = graphviz.Source(dot_graph)

#Membuat grafik menjadi pdf
graph.view()
```

Gambar 6. *Code Grafik Decision Tree*

Hasil akurasi dari penerapan algoritma C4.5 dapat dilihat sebagai berikut:

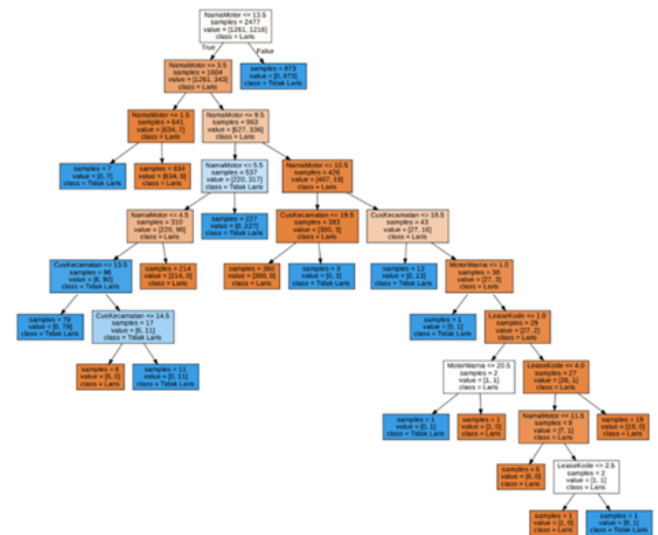
1. Data hasil proses *data mining* mendapatkan *accuracy* dan *confusion matrix* yang dapat dilihat pada gambar berikut ini:

Confusion Matrix  
 [[1286 0]  
 [ 1 1190]]  
 Tingkat Akurasi Algoritma C4.5  
 Akurasi :                      precision      recall      f1-score      support

0	1.00	1.00	1.00	1286
1	1.00	1.00	1.00	1191
accuracy			1.00	2477
macro avg	1.00	1.00	1.00	2477
weighted avg	1.00	1.00	1.00	2477

Tingkat Akurasi: 99 persen

Gambar 7. Hasil Akurasi



Gambar 8. Hasil Pohon Keputusan

2. Berdasarkan gambar diatas menjelaskan bahwa nilai akurasi dapat dijelaskan sebagai berikut:
  - a. Hasil akurasi pada penelitian ini sebesar 99%.
  - b. Prediksi Kurang Laris dengan true Kurang Laris sebanyak 1.286 data.
  - c. Prediksi Kurang Laris dengan true Laris sebanyak 1 data.
  - d. Prediksi laris dengan true Kurang Laris sebanyak 0 data.
  - e. Prediksi Laris dengan true Laris sebanyak 1.190 data.
3. Data hasil akurasi yang didapat menghasilkan pohon keputusan yang dapat dilihat pada gambar berikut ini:

Dari hasil pohon keputusan di atas, menghasilkan beberapa *rule* atau aturan yang terbentuk sebagai berikut:

- a. Jika (NamaMotor  $\leq 13.5$ , samples = 2477, values = [1261,1216], class = Laris) = False, maka keterangan “Tidak Laris”, Jika = True maka keterangan “Laris”.
- b. Jika (NamaMotor  $\leq 1.5$ , samples = 641, values = [634,7], class = Laris), dan samples = 7 maka keterangan “Tidak Laris”.
- c. Jika (NamaMotor  $\leq 1.5$ , samples = 641, values = [634,7], class = Laris), dan samples = 643 maka keterangan “Laris”.
- d. Jika (NamaMotor  $\leq 4.5$ , samples = 310, values = [220, 90], class = Laris), dan samples = 214 maka keterangan “Laris”.
- e. Jika (CusKecamatan  $\leq 13.5$ , samples = 96, values = [6,90], class = Tidak Laris), dan samples = 79 maka keterangan “Tidak Laris”.

## KESIMPULAN

Berikut ini adalah hasil kesimpulan yang dapat penulis sampaikan setelah dilakukan penelitian tentang penggunaan *data mining* dengan algoritma C4.5 *decision tree* pada prediksi produk laris sepeda motor Honda.

1. Penelitian ini telah menghasilkan klasifikasi penjualan sepeda motor Honda menggunakan bahasa pemrograman Python dengan alat bantu/tools yaitu Google Colab. Hasil akurasi pada penelitian ini menunjukkan angka 99%.
  2. Prediksi Kurang Laris dengan *true* Kurang Laris sebanyak 1.286 data. Prediksi Kurang Laris dengan *true* Laris sebanyak 1 data. Prediksi Laris dengan *true* Kurang Laris sebanyak 0 data. Dan Prediksi Laris dengan *true* Laris sebanyak 1.190 data.
  3. Variabel yang digunakan dalam penelitian ini yaitu CusKecamatan, GCust, NamaMotor, WarnaMotor, Segmen, LeaseKode, dan Keterangan sebagai variabel class. Kemudian data penelitian yang digunakan menggunakan data penjualan tahun 2020 dan data stok barang per bulan Desember 2020 pada CV Cendana Motor Cepiring Kendal.
  4. Dari pohon keputusan yang telah terbentuk, variabel tertinggi dalam memprediksi produk laris adalah variabel NamaMotor, artinya jenis motor sangat berpengaruh terhadap penjualan.
- DAFTAR PUSTAKA**
- [1] N. Azwanti, "Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada Pt. Capella Dinamik Nusantara Cabang Muka Kuning," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 13, no. 1, p. 33, 2018, doi: 10.30872/jim.v13i1.629.
  - [2] J. Eska, "Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5," vol. 2, 2018, doi: 10.31227/osf.io/x6svc.
  - [3] A. S. Sunge, H. Fidiawan, P. Studi, T. Informatika, S. Tinggi, and T. Pelita, "Data Mining, Penjualan Produk, Decision Tree, Algoritma C4.5 .," vol. 9, pp. 97–103, 2019.
  - [4] L. Elvitaria and M. Havenda, "Luluk Elvitaria, 2) Muhammad Havenda," vol. 2, no. 2, pp. 110–124, 2017.
  - [5] J. A. D. I. Swara, "Aplikasi Data Mining Menggunakan Metode Decision Tree Untuk Pemantauan Distribusi Penjualan Sepeda Motor Di Pd . Wijaya Abadi Bandung," pp. 1–8.
  - [6] D. A. Kurnia *et al.*, "KLASIFIKASI PRODUK PENJUALAN BROWNIES MENGGUNAKAN Jurnal Sistem Informasi dan Manajemen," vol. 9, no. 3, 2021.
  - [7] I. Sutoyo, "Implementasi Algoritma Decision Tree Untuk Klasifikasi Data Peserta Didik," *J. Pilar Nusa Mandiri*, vol. 14, no. 2, p. 217, 2018, doi: 10.33480/pilar.v14i2.926.
  - [8] A. F. Izzulhaq and Sulastri, "Klasifikasi Penjualan Aplikasi Android," *Proceeding SENDIU*, vol. 72, no. 6, pp. 978–979, 2020.
  - [9] M. Algoritma, "Klasifikasi Diagnosa Penyakit Paru-Paru Pada Klinik Raditya Medical," vol. 12, no. 1, pp. 1894–1905, 2020.
  - [10] A. K. Lalo, P. Batarius, and Y. C. H. Siki, "Implementasi Algoritma C4.5 Untuk Klasifikasi Penjualan Barang di Swalayan Dutalia," *J. Tek. Inform. UNIKA St. Thomas*, no. June, pp. 1–12, 2021, doi: 10.54367/jtiust.v6i1.1089.
  - [11] N. Dicky and W. N. Gunadi, "Algoritma Data Mining Dan Pengujian," Penerbit CV Budi Utama, Yogyakarta. 2019.
  - [12] A. Muhammad and N. Muhammad, "Data Mining – Algoritma dan Implementasi," Penerbit Andi, 2020.