

**ANALISA PENJUALAN ROKOK DENGAN  
METODE KLASIFIKASI MENGGUNAKAN  
FORMULA ALGORITMA C4.5 PADA  
CV JAYA BERKAH MAS**



**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat kelulusan Program Strata Satu

**SEBASTIANUS ADITYA NARENDRA      19236055**

**PRADITA ALDI SETIAWAN            19236049**

**Program Studi Sistem Informasi Kampus Kota Jakarta**

**Fakultas Teknik dan Informatika**

**Universitas Bina Sarana Informatika**

**2024**

# SURAT PERNYATAAN KEASLIAN SKRIPSI

## SURAT PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan dibawah ini:

Nama : Sebastianus Aditya Narendra

NIM : 19236055

Jenjang : Strata Satu (S1)

Program Studi : Sistem Informasi

Fakultas : Teknologi Informasi

Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa Skripsi yang telah saya buat dengan judul:

**“Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV Jaya Berkah Mas”**, adalah asli (orisinil) atau tidak plagiat (menjiplak) dan belum pernah diterbitkan/dipublikasikan dimanapun dan dalam bentuk apapun.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya dan tanpa ada paksaan dari pihak manapun. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa Skripsi yang telah saya buat adalah hasil karya seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari **Universitas Bina Sarana Informatika** dicabut/dibatalkan.

Dibuat di : Jakarta

Pada tanggal : 04 Juli 2024

Yang menyatakan,

Anggota :

1. Sebastianus Aditya Narendra

2. Pradita Aldi Setiawan



**Sebastianus Aditya Narendra**

# SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

## SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan dibawah ini:

Nama : Sebastianus Aditya Narendra  
NIM : 19236055  
Jenjang : Strata Satu (S1)  
Program Studi : Sistem Informasi  
Fakultas : Teknologi Informasi  
Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa seluruh data, informasi, interpretasi serta pernyataan yang terdapat dalam karya ilmiah Penulis dengan judul “**Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV Jaya Berkah Mas**” ini, kecuali yang disebutkan sumbernya adalah hasil pengamatan, penelitian, pengelolaan, serta pemikiran saya.

Penulis menyetujui untuk memberikan ijin kepada pihak Universitas Bina Sarana Informatika untuk mendokumentasikan karya ilmiah saya tersebut secara internal dan terbatas, serta tidak untuk mengunggah karya ilmiah Penulis pada *repository* Universitas Bina Sarana Informatika.

Penulis bersedia untuk bertanggung jawab secara pribadi, tanpa melibatkan pihak Universitas Bina Sarana Informatika, atas materi/isi karya ilmiah tersebut, termasuk bertanggung jawab atas dampak atau kerugian yang timbul dalam bentuk akibat Tindakan yang berkaitan dengan data, informasi, interpretasi serta pernyataan yang terdapat pada karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta  
Pada tanggal: 04 Juli 2024  
Yang menyatakan,



Sebastianus Aditya Narendra  
NIM. 19236055

# SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

## SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan dibawah ini:

Nama : Pradita Aldi Setiawan  
NIM : 19236049  
Jenjang : Strata Satu (S1)  
Program Studi : Sistem Informasi  
Fakultas : Teknologi Informasi  
Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa seluruh data, informasi, interpretasi serta pernyataan yang terdapat dalam karya ilmiah Penulis dengan judul "**Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV Jaya Berkah Mas**" ini, kecuali yang disebutkan sumbernya adalah hasil pengamatan, penelitian, pengelolaan, serta pemikiran saya.

Penulis menyetujui untuk memberikan ijin kepada pihak Universitas Bina Sarana Informatika untuk mendokumentasikan karya ilmiah saya tersebut secara internal dan terbatas, serta tidak untuk mengunggah karya ilmiah Penulis pada *repository* Universitas Bina Sarana Informatika.

Penulis bersedia untuk bertanggung jawab secara pribadi, tanpa melibatkan pihak Universitas Bina Sarana Informatika, atas materi/isi karya ilmiah tersebut, termasuk bertanggung jawab atas dampak atau kerugian yang timbul dalam bentuk akibat Tindakan yang berkaitan dengan data, informasi, interpretasi serta pernyataan yang terdapat pada karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta  
Pada tanggal: 04 Juli 2024  
Yang menyatakan,



Pradita Aldi Setiawan  
NIM. 19236049

# PERSETUJUAN DAN PENGESAHAN SKRIPSI

## PERSETUJUAN DAN PENGESAHAN SKRIPSI

Skripsi ini diajukan oleh:

Nama : Sebastianus Aditya Narendra  
NIM : 19236055  
Jenjang : Sarjana (S1)  
Program Studi : Sistem Informasi  
Fakultas : Teknik dan Informatika  
Perguruan Tinggi : Universitas Bina Sarana Informatika  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV. Jaya Berkah Mas

Telah dipertahankan pada periode 2024-1 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Komputer (S.Kom) pada Program Sarjana (S1) Program Studi Sistem Informasi di Universitas Bina Sarana Informatika.

Jakarta, 02 Agustus 2024

### PEMBIMBING SKRIPSI

Pembimbing I : Syaiful Anwar, M.Kom.

Pembimbing II : Mohammad Haddiel Fuad, S.Kom.,  
M.T.I.

### DEWAN PENGUJI

Penguji I : Indarti, M.Kom.

Penguji II : Rudianto, M.Kom.

# PERSETUJUAN DAN PENGESAHAN SKRIPSI

## PERSETUJUAN DAN PENGESAHAN SKRIPSI

Skripsi ini diajukan oleh:

Nama : Pradita Aldi Setiawan  
NIM : 19236049  
Jenjang : Sarjana (S1)  
Program Studi : Sistem Informasi  
Fakultas : Teknik dan Informatika  
Perguruan Tinggi : Universitas Bina Sarana Informatika  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV. Jaya Berkah Mas

Telah dipertahankan pada periode 2024-1 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Komputer (S.Kom) pada Program Sarjana (S1) Program Studi Sistem Informasi di Universitas Bina Sarana Informatika.

Jakarta, 02 Agustus 2024

### PEMBIMBING SKRIPSI

Pembimbing I : Syaiful Anwar, M.Kom.

Pembimbing II : Mohammad Haddiel Fuad, S.Kom.,  
M.T.I.

### DEWAN PENGUJI

Penguji I : Indarti, M.Kom

Penguji II : Rudianto, M.Kom.

## PEDOMAN PENGGUNAAN HAK CIPTA

Skripsi sarjana yang berjudul “**Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV Jaya Berkah Mas**” adalah hasil karya tulis asli Sebastianus Aditya Narendra dan Pradita Aldi Setiawan bukan hasil terbitan sehingga peredaran karya tulis hanya berlaku di lingkungan akademik saja, serta memiliki hak cipta. Oleh karena itu, dilarang keras untuk menggandakan baik sebagian maupun seluruhnya karya tulis ini, tanpa seizin penulis.

Referensi kepustakaan diperkenankan untuk dicatat tetapi pengutipan atau peringkasan isi tulisan hanya dapat dilakukan dengan seizin penulis dan disertai ketentuan pengutipan secara ilmiah dengan menyebutkan sumbernya.

Untuk keperluan perizinan pada pemilik dapat menghubungi informasi yang tertera di bawah ini:

Nama : Sebastianus Aditya Narendra

Nama : Pradita Aldi Setiawan

Alamat: Jl. Pd. Pakulonan Blok H6 No.3, Pakualam, Kec. Serpong Utara,  
Kota Tangerang Selatan, Banten 15320

No. Hp: 0812-2994-2990

e-mail : [adityanarendra021218@gmail.com](mailto:adityanarendra021218@gmail.com)



## LEMBAR KONSULTASI SKRIPSI

UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19236055  
Nama Lengkap : Sebastianus Aditya Narendra  
Dosen Pembimbing I : Syaiful Anwar, M.Kom  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing I
1	08 April 2024	Pengajuan Judul (online)	- <i>[Signature]</i>
2	22 April 2024	Pengajuan BAB I dan BAB II (online)	- <i>[Signature]</i> <i>[Signature]</i>
3	04 Mei 2024	Revisi BAB I (online)	- <i>[Signature]</i>
4	18 Mei 2024	Konsultasi BAB III (online)	- <i>[Signature]</i> <i>[Signature]</i>
5	31 Mei 2024	Pengajuan BAB III	- <i>[Signature]</i> <i>[Signature]</i>
6	26 Juni 2024	Revisi Data dan BAB III	- <i>[Signature]</i> <i>[Signature]</i>
7	27 Juni 2024	Penyesuaian Data dan BAB III	- <i>[Signature]</i> <i>[Signature]</i>
8	03 Juli 2024	ACC Keseluruhan	- <i>[Signature]</i> <i>[Signature]</i>

Catatan untuk Dosen Pembimbing.

Bimbingan Skripsi

Dimulai pada tanggal : 03 April 2024

Diakhiri pada tanggal : 03 Juli 2024

Jumlah pertemuan bimbingan : 8 (Delapan)

Disetujui oleh,  
Dosen Pembimbing I

(Syaiful Anwar, M.Kom)



## LEMBAR KONSULTASI SKRIPSI

UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19236055  
Nama Lengkap : Sebastianus Aditya Narendra  
Dosen Pembimbing I : Mohammad Haddiel Fuad, MTI  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing II
1	06 Mei 2024	Pengajuan dan ACC Judul (online)	<i>Fuad</i>
2	18 Mei 2024	Revisi BAB I dan BAB II (online)	<i>Fuad</i>
3	27 Juni 2024	Revisi Format Penulisan	<i>Fuad</i>
4	28 Juni 2024	Konsultasi BAB III (online)	<i>Fuad</i>
5	01 Juli 2024	Pengajuan BAB III	<i>Fuad</i>
6	02 Juli 2024	Revisi Data dan BAB III	<i>Fuad</i>
7	03 Juli 2024	Penyesuaian Data dan BAB III	<i>Fuad</i>
8	03 Juli 2024	ACC Keseluruhan	<i>Fuad</i>

Catatan untuk Dosen Pembimbing.

Bimbingan Skripsi

Dimulai pada tanggal : 06 Mei 2024

Diakhiri pada tanggal : 03 Juli 2024

Jumlah pertemuan bimbingan : 8 (Delapan)

Disetujui oleh,  
Dosen Pembimbing II

(Mohammad Haddiel Fuad, MTI)



## LEMBAR KONSULTASI SKRIPSI

UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19236049  
Nama Lengkap : Pradita Aldi Setiawan  
Dosen Pembimbing I : Syaiful Anwar, M.Kom  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing I
1	08 April 2024	Pengajuan Judul (online)	
2	22 April 2024	Pengajuan BAB I dan BAB II (online)	
3	04 Mei 2024	Revisi BAB I (online)	
4	18 Mei 2024	Konsultasi BAB III (online)	
5	31 Mei 2024	Pengajuan BAB III	
6	26 Juni 2024	Revisi Data dan BAB III	
7	27 Juni 2024	Penyesuaian Data dan BAB III	
8	03 Juli 2024	ACC Keseluruhan	

Catatan untuk Dosen Pembimbing.

Bimbingan Skripsi

Dimulai pada tanggal : 03 April 2024

Diakhiri pada tanggal : 03 Juli 2024

Jumlah pertemuan bimbingan : 8 (Delapan)

Disetujui oleh,  
Dosen Pembimbing I

(Syaiful Anwar, M.Kom)



## LEMBAR KONSULTASI SKRIPSI

UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19236049  
Nama Lengkap : Pradita Aldi Setiawan  
Dosen Pembimbing I : Mohammad Haddiel Fuad, MTI  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing II
1	06 Mei 2024	Pengajuan dan ACC Judul (online)	<i>[Signature]</i>
2	18 Mei 2024	Revisi BAB I dan BAB II (online)	<i>[Signature]</i>
3	27 Juni 2024	Revisi Format Penulisan	<i>[Signature]</i>
4	28 Juni 2024	Konsultasi BAB III (online)	<i>[Signature]</i>
5	01 Juli 2024	Pengajuan BAB III	<i>[Signature]</i>
6	02 Juli 2024	Revisi Data dan BAB III	<i>[Signature]</i>
7	03 Juli 2024	Penyesuaian Data dan BAB III	<i>[Signature]</i>
8	03 Juli 2024	ACC Keseluruhan	<i>[Signature]</i>

Catatan untuk Dosen Pembimbing.

Bimbingan Skripsi

Dimulai pada tanggal : 06 Mei 2024

Diakhiri pada tanggal : 03 Juli 2024

Jumlah pertemuan bimbingan : 8 (Delapan)

Disetujui oleh,  
Dosen Pembimbing II

*[Signature]*  
(Mohammad Haddiel Fuad, MTI)

## PERSEMBAHAN

*Ilmu adalah yang memberikan manfaat, bukan yang sekedar hanya dihafal.*

(Imam Syafi'i)

Segala puji bagi Allah S.W.T yang maha menguasai semesta alam beserta isinya, skripsi ini penulis persembahkan untuk:

1. Orangtua tercinta yang telah melahirkan, merawat, mendidik, memotivasi, serta memberi apa yang terbaik untuk menuju kesuksesan dimasa yang akan datang.
2. Teman-teman seperjuangan yang telah memberikan bantuan material maupun spiritual demi terselesaikannya skripsi ini.
3. Teman-teman kelompok yang selalu gotong royong dalam menyelesaikan masalah satu sama lain.



## KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas segala nikmat, rahmat dan karunianya, sehingga pada akhirnya penulis dapat menyelesaikan tugas ini dengan baik. Skripsi pada Program Sarjana (S1) ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul Skripsi yang penulis ambil sebagai berikut, “**Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV Jaya Berkah Mas**”.

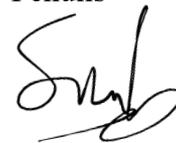
Tujuan penulisan Skripsi pada Program Sarjana (S1) ini dibuat sebagai salah satu syarat kelulusan Program Sarjana Universitas Bina Sarana Informatika. Sebagai bahan penulisan diambil berdasarkan hasil penelitian (eksperimen), observasi dan beberapa sumber literatur yang mendukung penulisan ini. Penulis menyadari bahwa tanpa bimbingan dan dorongan dari semua pihak, maka penulisan Skripsi ini tidak akan berjalan lancar. Oleh karena itu pada kesempatan ini, izinkanlah penulis menyampaikan ucapan terimakasih kepada:

1. Rektor Universitas Bina Sarana Informatika.
2. Dekan Fakultas Teknologi Informasi Universitas Bina Sarana Informatika.
3. Ketua Program Studi Ilmu Komputer Universitas Bina Sarana Informatika Kampus Sukabumi.
4. Bapak Syaiful Anwar, M.Kom. dan Bapak Mohammad Haddiel Fuad, MTI. selaku Dosen Pembimbing Skripsi.
5. Staff / karyawan / dosen di lingkungan Universitas Bina Sarana Informatika Kampus Kramat 98.
6. Bapak Desta Pratama selaku Direktur CV Jaya Berkah Mas.
7. Bapak Mugiyono selaku Pemilik Kantor CV Jaya Berkah Mas.
8. Bapak Adi Wardoyo selaku Kepala Gudang CV Jaya Berkah Mas.
9. Staf / karyawan di lingkungan CV Jaya Berkah Mas.
10. Orang tua tercinta yang telah memberikan dukungan moral maupun spiritual.
11. Rekan-rekan mahasiswa kelas 19.8AH.07 dan 19.8AG.07.

Serta semua pihak yang terlalu banyak untuk disebut satu persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan Skripsi ini masih jauh dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan di masa yang akan datang.

Akhir kata semoga Skripsi ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 08 April 2024  
Penulis



**Sebastianus Aditya Narendra**

## ABSTRAK

**Sebastianus Aditya Narendra (19236055), Pradita Aldi Setiawan (19236049),  
Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula  
Algoritma C4.5 Pada CV Jaya Berkah Mas**

Penelitian ini bertujuan untuk menganalisis dan mengklasifikasikan data penjualan rokok menggunakan algoritma C4.5 pada CV Jaya Berkah Mas. Penelitian dilakukan dengan mengumpulkan data melalui observasi dan wawancara, serta melakukan pembersihan data untuk mengatasi inkonsistensi dan duplikasi. Algoritma C4.5 digunakan untuk menghasilkan pohon keputusan yang dapat memprediksi penjualan rokok terlaris. Hasil dari analisis menunjukkan bahwa metode C4.5 dapat digunakan untuk memberikan wawasan yang bermanfaat dalam menentukan strategi penjualan yang optimal. Penelitian ini memberikan kontribusi dalam bidang data mining dengan menerapkan algoritma C4.5 pada kasus nyata di industri rokok. Penelitian ini menghasilkan atribut baru sebagai output pada pohon keputusan seperti rasa, kemasan rokok, aroma dan hisapan rokok yang dihasilkan dari membandingkan penelitian terdahulu. Hasil nilai akurasi tertinggi perhitungan Algoritma C4.5 dengan data sebanyak 212 *entry* menunjukkan tingkat Akurasi 96.73%, Presisi 98.00%, *Recall* 85.17%, dan *Area Under Curve* (AUC) 0.970 yang termasuk kategori sangat baik.

**Kata Kunci : Penjualan rokok, Data Mining, Algoritma C4.5, Decision Tree, Rapidminer, Klasifikasi**



UNIVERSITAS

## **ABSTRACT**

***Sebastianus Aditya Narendra (19236055), Pradita Aldi Setiawan (19236049),  
Analysis of Cigarette Sales with Classification Methods Using the C4.5 Algorithm  
Formula at CV Jaya Berkah Mas***

*This study aims to analyze and classify cigarette sales data using the C4.5 algorithm at CV Jaya Berkah Mas. The research was conducted by collecting data through observation and interviews, as well as performing data cleaning to overcome inconsistencies and duplication. The C4.5 algorithm is used to generate a decision tree that can predict the best-selling cigarette sales. The results of the analysis show that the C4.5 method can be used to provide useful insights in determining the optimal sales strategy. This research contributes to the field of data mining by applying the C4.5 algorithm to real cases in the cigarette industry. This research produces new attributes as output on the decision tree such as flavor, cigarette packaging, aroma and cigarette suction resulting from comparing previous research. The results of the highest accuracy value of the C4.5 algorithm calculation with data as many as 212 entries show the level of Accuracy 96.73%, Precision 98.00%, Recall 85.17%, and Area Under Curve (AUC) 0.970 which is included in the excellent category.*

***Keywords: Cigarette sales, Data Mining, C4.5 Algorithm, Decision Tree, Rapidminer, classification***



UNIVERSITAS

## DAFTAR ISI

<b>LEMBAR JUDUL SKRIPSI</b> .....	<b>i</b>
<b>SURAT PERNYATAAN KEASLIAN SKRIPSI</b> .....	<b>ii</b>
<b>SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS</b> .....	<b>iii</b>
<b>PERSETUJUAN DAN PENGESAHAN SKRIPSI</b> .....	<b>v</b>
<b>PEDOMAN PENGGUNAAN HAK CIPTA</b> .....	<b>vii</b>
<b>LEMBAR KONSULTASI SKRIPSI</b> .....	<b>viii</b>
<b>PERSEMBAHAN</b> .....	<b>xii</b>
<b>KATA PENGANTAR</b> .....	<b>xiii</b>
<b>ABSTRAK</b> .....	<b>xiv</b>
<b>ABSTRACT</b> .....	<b>xv</b>
<b>DAFTAR ISI</b> .....	<b>xvi</b>
<b>DAFTAR SIMBOL</b> .....	<b>xviii</b>
<b>DAFTAR GAMBAR</b> .....	<b>xx</b>
<b>DAFTAR TABEL</b> .....	<b>xxi</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xxii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Tujuan Penelitian.....	2
1.3. Penelitian Terdahulu.....	2
1.4. Gambaran Metode Usulan.....	4
1.5. <i>State Of The Art</i> (Kebaruan).....	6
<b>BAB II METODOLOGI PENELITIAN</b> .....	<b>7</b>
2.1. Tahapan Penelitian .....	7
2.2. Instrumen Penelitian.....	9
2.3. Metode Pengumpulan Data .....	9
2.4. Metode Analisis Data .....	10
2.4.1. <i>Data Mining</i> .....	10
2.4.2. Rapidminer .....	10
2.4.3. Klasifikasi .....	10
2.4.4. Algoritma C4.5.....	11

2.4.5. <i>Decision Tree</i> .....	12
<b>BAB III HASIL DAN PEMBAHASAN .....</b>	<b>13</b>
3.1. Pengumpulan Data .....	13
3.2. Tahap <i>Pre-Processing</i> .....	14
3.2.1. Seleksi Data.....	14
3.3. Pengolahan Data Menggunakan Algoritma C4.5.....	14
3.3.1. <i>Decision Tree</i> .....	14
3.3.2. Algoritma C4.5.....	40
<b>BAB IV KESIMPULAN .....</b>	<b>44</b>
4.1. Kesimpulan.....	44
<b>DAFTAR PUSTAKA.....</b>	<b>46</b>
<b>DAFTAR RIWAYAT HIDUP .....</b>	<b>48</b>
<b>LEMBAR KONSULTASI BIMBINGAN.....</b>	<b>50</b>
<b>SURAT KETERANGAN RISET .....</b>	<b>54</b>
<b>LAMPIRAN – LAMPIRAN.....</b>	<b>56</b>
<b>BUKTI HASIL PENGECEKAN PLAGIARISME.....</b>	<b>59</b>



UNIVERSITAS

## DAFTAR SIMBOL

### 1. Simbol UML (*Unified Modelling Language*)

#### a. Simbol *Activity Diagram*

	<p><i>STATUS AWAL/INITIAL</i></p> <p>Status awal aktivitas sistem, sebuah diagram aktivitas memiliki sebuah status awal.</p>
	<p><i>AKTIVITAS/ACTIVITY</i></p> <p>Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja.</p>
	<p><i>PERCABANGAN/DECISION</i></p> <p>Asosiasi percabangan dimana lebih dari satu aktivitas digabungkan menjadi satu.</p>
	<p><i>PENGGABUNGAN/JOIN</i></p> <p>Asosiasi penggabungan dimana lebih dari satu aktivitas lebih dari satu.</p>
	<p><i>STATUS AKHIR/FINAL</i></p> <p>Status akhir yang dilakukan sistem, sebuah diagram aktivitas memiliki sebuah status satu.</p>
	<p><i>SWIMLINE</i></p> <p>Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi.</p>

b. Simbol *Entropy*

Simbol	Arti
$S$	Himpunan Kasus
$n$	Jumlah partisi A
$p_i$	Proporsi $S_i$ terhadap $S$
$A$	Jumlah Partisi A
$ S_i $	Proporsi $S_i$ terhadap $S$
$ S $	Jumlah kasus dalam $S$
$Y$	Data dengan kelas yang belum diketahui
$X$	Hipotesis data $y$ merupakan suatu kelas spesifik
$(x y)$	Probabilitas hipotesis $x$ berdasarkan kondisi $y$ ( <i>posterior probability</i> )
$P(x)$	Probabilitas hipotesis $x$ ( <i>prior probability</i> )
$P(y x)$	Probabilitas $y$ berdasarkan kondisi pada hipotesis $x$
$P(y)$	Probabilitas $y$



## DAFTAR GAMBAR

<b>Gambar II. 1.</b> Tahapan Penelitian.....	7
<b>Gambar III. 1.</b> Data Penjualan Rokok.....	13
<b>Gambar III. 2.</b> Dataset yang telah di seleksi .....	14
<b>Gambar III. 3.</b> Proses <i>Decision Tree</i> .....	15
<b>Gambar III. 4.</b> Atribut Rasa Sebagai <i>root</i> pohon keputusan .....	29
<b>Gambar III. 5.</b> Atribut Kemasan Rokok sebagai <i>root</i> pohon keputusan .....	30
<b>Gambar III. 6.</b> Atribut Aroma sebagai <i>root</i> pohon keputusan .....	31
<b>Gambar III. 7.</b> Atribut Hisapan Rokok sebagai <i>root</i> pohon keputusan.....	33
<b>Gambar III. 8.</b> Dataset Penjualan Rokok .....	33
<b>Gambar III. 9.</b> Penambahan Label Pada Atribut Label .....	34
<b>Gambar III. 10.</b> Pengolahan <i>Decision Tree</i> Menggunakan Rapidminer .....	35
<b>Gambar III. 11.</b> Hasil <i>Decision Tree</i> pada Rapidminer .....	35
<b>Gambar III. 12.</b> Rule hasil algoritma <i>Decision Tree</i> .....	36
<b>Gambar III. 13.</b> Proses Pengolahan Akurasi Data pada Rapidminer .....	37
<b>Gambar III. 14.</b> Nilai <i>Accuracy</i> pada <i>Decision Tree</i> .....	38
<b>Gambar III. 15.</b> Nilai <i>Precision</i> pada <i>Decision Tree</i> .....	38
<b>Gambar III. 16.</b> Nilai <i>Recall</i> pada <i>Decision Tree</i> .....	38
<b>Gambar III. 17.</b> Nilai <i>Area Under Curve</i> (AUC) pada <i>Decision Tree</i> .....	39
<b>Gambar III. 18.</b> Dataset Penjualan Rokok .....	40
<b>Gambar III. 19.</b> Penambahan Label Pada Atribut Label .....	41
<b>Gambar III. 20.</b> Proses Pengolahan Akurasi Data pada Rapidminer .....	41
<b>Gambar III. 21.</b> Nilai <i>Accuracy</i> pada Algoritma C4.5 .....	42
<b>Gambar III. 22.</b> Nilai <i>Precision</i> pada Algoritma C4.5 .....	42
<b>Gambar III. 23.</b> Nilai <i>Recall</i> pada Algoritma C4.5 .....	42
<b>Gambar III. 24.</b> Nilai <i>Area Under Curve</i> (AUC) pada Algoritma C4.5.....	43

## DAFTAR TABEL

<b>Tabel III. 1.</b> Tabel Atribut Kategori Rokok.....	16
<b>Tabel III. 2.</b> Tabel Atribut Jenis Rokok.....	18
<b>Tabel III. 3.</b> Tabel Atribut Aroma.....	19
<b>Tabel III. 4.</b> Tabel Atribut Rasa.....	20
<b>Tabel III. 5.</b> Tabel Atribut Ukuran Batang Rokok.....	23
<b>Tabel III. 6.</b> Tabel Atribut Kepadatan Rokok.....	24
<b>Tabel III. 7.</b> Tabel Atribut Hisapan Rokok.....	25
<b>Tabel III. 8.</b> Tabel Atribut Kemasan Rokok.....	26
<b>Tabel III. 9.</b> Tabel Penentuan <i>Node</i> 1 Pohon Keputusan .....	27
<b>Tabel III. 10.</b> Tabel Penentuan <i>Node</i> 2 Pohon Keputusan .....	29
<b>Tabel III. 11.</b> Tabel Penentuan <i>Node</i> 3 Pohon Keputusan.....	30
<b>Tabel III. 12.</b> Tabel Penentuan <i>Node</i> 4 Pohon Keputusan .....	32



## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Produk Rokok .....	56
<b>Lampiran 2.</b> Jaya Berkah Mas Territory .....	57
<b>Lampiran 3.</b> Personil dan Jadwal Tim Jaya Berkah Purbalingga.....	58



# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Industri rokok merupakan industri yang dapat memberikan kontribusi terhadap PAD (Pendapatan Asli Daerah). Dengan semakin ketatnya persaingan di setiap industri, industri rokok juga harus mampu menentukan strategi pemasaran agar dapat bertahan dalam bisnisnya (Qomariah et al., 2024). JAYA BERKAH MAS, sebagai salah satu pengecer rokok di kawasan BarLingMasCaKeb (Banjarnegara, Purbalinga, Banyumas, Cilacap, Kebumen), berupaya mempertahankan dan meningkatkan posisi pasarnya dengan meningkatkan pelayanan dan penjualan yang efektif. Perusahaan menghadapi tantangan dalam menganalisis dan memahami pola konsumsi konsumen serta mengadaptasi strategi penjualan dan pemasaran terhadap berbagai kebutuhan dan preferensi konsumen.

Namun di era digital saat ini JAYA BERKAH MAS ditantang untuk memanfaatkan teknologi dan data untuk mendapatkan keunggulan kompetitif. Teknik pengumpulan data penting sebagai alat analisis data yang efektif untuk menganalisis data penjualan rokok dan mengidentifikasi faktor-faktor yang mempengaruhi penjualan. Oleh karena itu, permasalahan penelitian ini bertujuan untuk mengklasifikasikan penjualan rokok dengan menggunakan metode Algoritma C4.5 dengan atribut Merek, Kategori, Jenis, Aroma, Rasa dan atribut lainnya yang relevan tentang Rokok supaya akurasi yang tinggi.

Algoritma C4.5 merupakan perkembangan dari algoritma ID3 yang diciptakan oleh J. Ross Quinlan (Tukino, 2019). Algoritma ini merupakan bagian dari algoritma

pembelajaran pohon keputusan (*Decision Tree Learning*) yang dirancang untuk meningkatkan efisiensi dan efektivitas penjualan (Suriani, 2023). Selain itu, algoritma ini membantu memahami tren konsumen dan pasar dengan lebih baik, sehingga dapat menyusun strategi penjualan yang lebih efektif.

Dengan menggunakan metode Algoritma C4.5 diharapkan membantu perhitungan penjualan rokok penelitian ini di angkat dengan judul **“Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV Jaya Berkah Mas”**.

## **1.2. Tujuan Penelitian**

Tujuan dari penelitian ini yaitu untuk Menganalisis dan Mengklasifikasikan Data Penjualan Rokok serta sebagai salah satu syarat kelulusan Program Srata Satu (S1) Jurusan Sistem Informasi di Universitas Bina Sarana Informatika Kramat 98 Jakarta.

## **1.3. Penelitian Terdahulu**

Penelitian terdahulu yang digunakan adalah penelitian yang serupa dengan penelitian ini dari lima tahun sebelumnya:

1. Penelitian terdahulu yang dilakukan oleh (Leonardi et al., 2021): “Prediksi Penjualan Produk Rokok Pada PT. Indomarco Prismatama Menggunakan Algoritma C4.5”, penelitian ini menggunakan 58 data sebagai data pelatihan dan 37 data sebagai data pengujian, dengan total keseluruhan 95 data. Evaluasi hasil penerapan algoritma C4.5 dengan menggunakan 37 data pengujian menunjukkan akurasi sebesar 92,11% dan AUC sebesar 0,878, yang menunjukkan tingkat akurasi dalam kategori *Good Classification*.
2. Penelitian terdahulu yang dilakukan oleh (Triawan et al., 2024): “Penerapan Data Mining untuk Prediksi Penjualan Spanduk Menggunakan Algoritma C4.5”,

Penelitian ini mengumpulkan data penjualan spanduk dari Januari hingga Juni 2023 dan menggunakan algoritma C4.5. Melalui proses data mining dengan aplikasi Rapidminer 5, dihasilkan pengetahuan baru berupa enam aturan dengan tingkat kinerja algoritma C4.5 sebesar 84%. Aturan-aturan ini dapat digunakan oleh manajemen perusahaan sebagai dasar untuk memprediksi pencapaian target penjualan.

3. Penelitian terdahulu yang dilakukan oleh (Musa et al., 2024): “Penerapan Data Mining Untuk Klasifikasi Data Penjualan Pakan Ternak Terlaris Dengan Algoritma C4.5”, Penelitian ini menghitung *Entropy* dan *Gain* sampai tahap akhir, yaitu pembentukan pohon keputusan berdasarkan perhitungan tersebut. Hasil penelitian menunjukkan bahwa kategori pakan memiliki *gain* tertinggi sebesar 0.306739968 dan *entropy* pakan ayam pedaging sebesar 0.99107606. Hal ini mengindikasikan bahwa pakan ayam pedaging merupakan produk terlaris berdasarkan hasil analisis data.
4. Penelitian terdahulu yang dilakukan oleh (Wahyudi, 2023): “Penerapan Data Mining Dengan Algoritma C4.5 Dalam Prediksi Penjualan Buku”, Penelitian ini bertujuan untuk membantu penerbit memprediksi penjualan buku dan menyiapkan stok berdasarkan prediksi peningkatan penjualan. Dari penelitian ini, disimpulkan bahwa dengan menggunakan algoritma C4.5, penjualan buku dapat diprediksi dengan akurasi yang tinggi. Hasil pengukuran akurasi data menunjukkan bahwa dari 90 data pelatihan, *Confusion Matrix* menghasilkan nilai akurasi 100%, dan data pengujian juga mencapai tingkat akurasi 100%.
5. Penelitian terdahulu yang dilakukan oleh (Ridho & Rusda, 2024): “Analisis Preferensi Konsumen dalam Memilih Produk Hortikultura Menggunakan Metode Algoritma C45 dan Naive Bayes’, Penelitian ini bertujuan untuk analisis

preferensi konsumen terhadap kombinasi atribut tomat dan cabai rawit di Pasar Keramat Baamang Sampit. Dua metode algoritma yang digunakan adalah C4.5 dan Naive Bayes dengan software RapidMiner untuk klasifikasi. Hasil analisis menunjukkan bahwa algoritma C4.5 memiliki tingkat akurasi 60,00% hingga 63,06% untuk pemilihan tomat, sementara *Naive Bayes* memiliki akurasi 45,00% hingga 49,05%. Dalam pemilihan cabai rawit, algoritma C4.5 memiliki akurasi 89,00% hingga 89,05%, dan Naive Bayes memiliki akurasi 89,05% hingga 90,00%. Kesimpulannya, algoritma C4.5 lebih akurat untuk pemilihan tomat, sedangkan *Naive Bayes* lebih akurat untuk pemilihan cabai rawit.

6. Penelitian terdahulu yang dilakukan oleh (Marliani et al., 2023): “Analisa Standar Produk Sabun Pembersih Baju Menggunakan Algoritma C4.5”, Penelitian ini bertujuan untuk menganalisis penggunaan algoritma C4.5 untuk klasifikasi preferensi sabun pembersih pakaian, khususnya sabun transparan. Data dikumpulkan melalui 200 kuesioner Google Form terkait standar produk sabun pembersih pakaian. Proses algoritma C4.5 diimplementasikan menggunakan RapidMiner. Hasil penelitian menunjukkan akurasi sebesar 81,52%, dengan rincian: 264 data prediksi sesuai dan benar, 48 data prediksi tidak sesuai dan benar, 12 data prediksi tidak sesuai dan salah, dan 0 data prediksi sesuai dan salah. Nilai AUC yang diperoleh yaitu 0,559. Kesimpulannya, algoritma C4.5 menunjukkan kinerja akurasi yang baik dalam klasifikasi preferensi sabun pembersih pakaian.

#### **1.4. Gambaran Metode Usulan**

Pada penelitian ini penulis menerapkan metode usulan tentang "Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas" sebagai berikut:

## 1. Pengumpulan Data

### a. Data Penjualan

Mengumpulkan data penjualan rokok dari CV. JAYA BERKAH MAS, termasuk detail seperti tanggal penjualan, status penjualan, kode sales, nama sales, alamat pelanggan, produk rokok, total kuantitas, dan informasi lain yang relevan.

### b. Data Rokok

Mengumpulkan data rokok, seperti preferensi merek, ukuran rokok, rasa rokok, aroma rokok dan jumlah rokok setiap *Pieces*.

## 2. Analisis Data Menggunakan metode Algoritma C4.5

### a. Persiapan Data

Pemilihan Atribut, memilih atribut yang akan digunakan sebagai input dan output dari algoritma. Atribut *input* merupakan variabel yang mempengaruhi hasil, Jenis rokok, harga, jumlah rokok, dan rasa. Atribut *output* merupakan variabel yang ingin dihitung, seperti harga rokok. Pengkodean Data, Mengubah data kategorikal ke format numerik yang dapat diuraikan oleh algoritma C4.5.

### b. Pembuatan Model

Pembuatan Pohon Keputusan, menggunakan algoritma C4.5 untuk membuat pohon keputusan berdasarkan data yang telah dipersiapkan. Algoritma ini akan menganalisis data dan menghasilkan pohon keputusan yang mencerminkan hubungan antara atribut *input* dan *output*.

Pemilihan Aturan: Dari pohon keputusan yang dibuat, aturan-aturan dihasilkan yang mencerminkan hubungan antara atribut *input* dan *output*.

Aturan ini dapat digunakan untuk memprediksi hasil berdasarkan variabel input.

### c. Evaluasi Model

Penerapan Aturan, menggunakan aturan yang dihasilkan dari model untuk membuat keputusan dalam konteks penjualan rokok. Aturan ini dapat digunakan untuk menentukan strategi penjualan dan pemasaran yang lebih efektif. Pemantauan dan Penyesuaian, melakukan pemantauan dan penyesuaian model berdasarkan hasil dan *feedback* untuk memastikan model tetap relevan dan efektif.

### 3. Pengembangan Penjualan

Mengembangkan penjualan yang berfokus pada data rokok yang telah diidentifikasi, dengan mempertimbangkan aroma rokok, hisapan rokok, merek rokok, jenis rokok, jumlah rokok, dan rasa rokok.

#### 1.5. *State Of The Art* (Kebaruan)

Berdasarkan penjelasan yang sudah diuraikan diatas, metode algoritma C4.5 dapat memperhitungkan penjualan rokok dengan akurasi yang tinggi. Dengan menggunakan algoritma C4.5 penelitian ini, diharapkan menghasilkan akurasi yang lebih besar dari penelitian sebelumnya yang dilakukan oleh (Leonardi et al., 2021): “Prediksi Penjualan Produk Rokok Pada PT. Indomarco Prismatama Menggunakan Algoritma C4.5”.

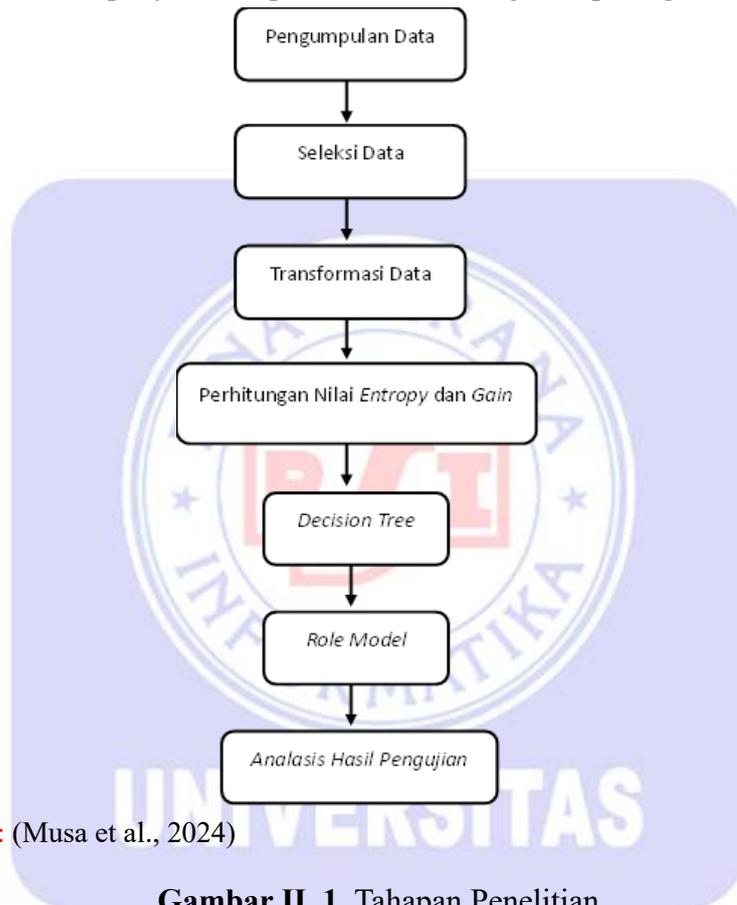
Untuk kebaruan dari penelitian yang kami lakukan yaitu menambahkan atribut yang tidak ada pada penelitian sebelumnya. Atribut yang kami tambahkan seperti jenis rokok, aroma, rasa, kepadatan rokok, hisapan rokok, dan kemasan rokok yang didapatkan dari hasil membandingkan penelitian terdahulu yang dilakukan oleh (Marliani et al., 2023) dan (Ridho & Rusda, 2024) dengan penelitian kami.

## BAB II

### METODOLOGI PENELITIAN

#### 2.1. Tahapan Penelitian

Tahapan dalam penyusunan penelitian ini ditunjukkan pada gambar II.1



Sumber: (Musa et al., 2024)

**Gambar II. 1.** Tahapan Penelitian

Adapun tahapan penelitian yang dilakukan oleh penulis dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan Data

Pengumpulan data dilakukan dengan observasi langsung serta melakukan wawancara dengan Kepala Gudang CV Jaya Berkah Mas, Jl. Mustari No.17,

RT.03/RW.09, Dusun 4, Penaruban, Kec. Kaligondang, Kabupaten Purbalingga, Jawa Tengah 53331.

## 2. Seleksi Data

Seleksi data untuk memilih data yang sesuai dengan kebutuhan untuk memperoleh pengetahuan. Data hasil seleksi digunakan dalam proses data mining, untuk memilih semua data yang diperlukan untuk diproses lebih lanjut. Data disimpan dalam file yang terpisah dari database produksi agar mudah digunakan nanti (Shaumi et al., 2022).

## 3. Transformasi Data

Transformasi data Disebut juga integrasi data. Fase ini mengubah data yang dipilih ke dalam format yang sesuai untuk pemrosesan semantik dengan menormalkan dan menggabungkannya (Zai, 2022).

## 4. Perhitungan Nilai *Entropy* dan *Gain*

Hitung entropi seluruh atribut/variabel, entropi persamaan (1), dan *information gain* persamaan (2) untuk mencari *information gain* tertinggi. Ini akan digunakan sebagai simpul akar ketika membangun pohon keputusan.

## 5. *Decision Tree*

Pohon keputusan merupakan hasil proses penghitungan entropi dan perolehan informasi setelah setiap atribut pohon mempunyai kelas dan proses penghitungan tersebut diulangi hingga tidak dapat dilakukan lagi.

## 6. *Role Model*

*Role Model*/Model aturan menggambarkan pohon keputusan.

## 7. Analisis Hasil Pengujian

Analisis hasil pengujian dalam *data mining* merupakan proses mengevaluasi model untuk menilai kinerja prediktif atau klasifikasinya. Pertama, bagi data

menjadi set pelatihan dan set pengujian. Metode evaluasi yang umum digunakan meliputi validasi silang, matriks konfusi (untuk menghitung presisi, presisi, perolehan, dan skor F1), serta kurva ROC dan AUC untuk mengukur performa model secara keseluruhan. Masalah regresi menggunakan metrik seperti *mean absolute error* (MAE) dan *mean squared error* (MSE). Hasil evaluasi kemudian diinterpretasikan untuk menentukan seberapa baik kinerja model.

## 2.2. Instrumen Penelitian

“Instrumen penelitian merupakan pedoman tertulis yang mencakup wawancara, pengamatan, dan pertanyaan yang disiapkan untuk memperoleh informasi.” (Ovan & Saputra, 2020).

## 2.3. Metode Pengumpulan Data

Dalam penelitian ini peneliti menggunakan 3 pengumpulan data sebagai berikut:

### 1. Observasi

Observasi merupakan teknik yang sangat umum digunakan dalam penelitian kuantitatif. Dalam observasi ini penulis melakukan riset di CV Jaya Berkah Mas untuk memperoleh data penelitian. Adapun lokasi observasi pada penelitian ini yaitu di Jl. Mustari No.17, RT.03/RW.09, Dusun 4, Penaruban, Kec. Kaligondang, Kabupaten Purbalingga, Jawa Tengah 53331.

### 2. Wawancara

Wawancara merupakan salah satu teknik utama untuk mengumpulkan data. Percakapan dalam wawancara terjadi antara dua orang atau lebih antara narasumber dan pewawancara. Dalam wawancara ini peneliti mewawancarai kepala staff gudang CV Jaya Berkah Mas.

### 3. Dokumentasi

Dokumentasi merupakan catatan otentik atau dokumen asli yang dapat digunakan sebagai bukti dalam masalah hukum. Dokumentasi dilakukan untuk mencatat, mendokumentasikan, dan menyimpan informasi yang telah diperoleh bisa berbentuk tulisan, gambar/foto dan sebagainya.

## **2.4. Metode Analisis Data**

Untuk mencapai tujuan penelitian maka analisa yang digunakan adalah analisis data kuantitatif. Analisis data kuantitatif merupakan suatu analisis data yang dipergunakan apabila kesimpulan-kesimpulan yang diperoleh berupa angka-angka dan juga dalam perhitungan dipergunakan rumus yang ada hubungannya dengan analisis penulisan (Umrati & Wijaya, 2020). Dalam hal ini akan dipergunakan analisis menggunakan metode Algoritma C4.5.

### **2.4.1. Data Mining**

Data Mining proses menggali informasi dan pola yang bermanfaat dari data yang sangat besar. Proses data *mining* meliputi pengumpulan data, ekstraksi data, analisis data, dan statistik data. Proses ini juga dikenal sebagai *knowledge discovery*, *knowledge extraction*, *data/pattern analysis*, dan *information harvesting*. Keempat proses dalam data mining ini menghasilkan model atau pengetahuan yang sangat berguna (Arhami & Nasir, 2020).

### **2.4.2. Rapidminer**

RapidMiner merupakan perangkat lunak sumber terbuka (open source) yang digunakan untuk analisis data mining dan prediksi. RapidMiner menawarkan berbagai teknik deskriptif dan prediktif untuk memberikan wawasan kepada pengguna, membantu mereka membuat keputusan yang optimal (Pratiwi et al., 2020).

### **2.4.3. Klasifikasi**

Febriani dan Sulistiani mengemukakan “Klasifikasi merupakan proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan memperkirakan kelas dari suatu objek yang labelnya belum diketahui. Model tersebut dapat berupa aturan “jika maka”, pohon keputusan, atau formula matematis” (Anggi Trifani et al., 2022).

#### 2.4.4. Algoritma C4.5

Algoritma C4.5 program yang memproses sekumpulan data untuk menghasilkan pohon keputusan sebagai *output*. Pohon keputusan yang dihasilkan kemudian divalidasi menggunakan data uji berlabel yang belum terlihat untuk menghitung generalisasinya (Sugara et al., 2019).

Untuk menyelesaikan kasus pada Algoritma C4.5 supaya lebih mudah dipahami prosesnya adalah sebagai berikut:

1. Mencari nilai *Gain*

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

Dimana keterangan dari rumus tersebut adalah:

S = Himpunan Kasus

A = atribut

n = jumlah partisi

pi = porsi dari si kepada s

|si| = jumlah kasus pada nilai atribut/kriteria

|S| = jumlah kasus atau total kasus

2. Mencari nilai *Entropy*

$$Entropy(S) = \sum_{i=1}^n - p_i * \text{Log}_2 p_i \dots \dots (2)$$

Keterangan:

S : Himpunan Kasus

N: Jumlah Partisi S

Pi: Proporsi dari Si terhadap S.

#### **2.4.5. *Decision Tree***

*Decision Tree* (Pohon Keputusan) merupakan salah satu metode klasifikasi yang paling populer karena mudah diinterpretasikan oleh manusia. *Decision Tree* merupakan model prediksi yang menggunakan struktur pohon atau hierarki. Konsep dasar dari *Decision Tree* yaitu mengubah data menjadi pohon keputusan dan aturan-aturan keputusan. Manfaat utama dari penggunaan *Decision Tree* adalah kemampuannya untuk memecah proses pengambilan keputusan yang kompleks menjadi lebih sederhana, sehingga mempermudah dalam menginterpretasikan solusi dari permasalahan (Fauziningrum & Suryaningsih, 2021).

## BAB III

### HASIL DAN PEMBAHASAN

#### 3.1. Pengumpulan Data

Riset yang telah dilakukan di CV Jaya Berkah Mas pada tanggal 10 April 2024 dan melakukan wawancara dengan mengajukan sebuah pertanyaan kepada Bapak Adi Wardoyo selaku Kepala Gudang CV Jaya Berkah Mas untuk mengetahui data-data yang akan menjadi perhitungan penjualan rokok menggunakan Algoritma C4.5 terhadap penelitian ini. Data yang diambil dari penjualan rokok di CV Jaya Berkah Mas per-tanggal 23 dan 24 April 2024. Memiliki data sebanyak 212 *Entry*.

NAMA PRODUK	Unit	Kategori Rokok	Jenis Rokok	Aroma	Rasa	Ukuran Batang Rokok	Kepadatan Rokok	Hisapan Rokok	Kemasan Rokok	Label
CENGKEH ABADI	6 PCS	SKT	Kretek	Sangat Wangi	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
CENGKEH ABADI	10 PCS	SKT	Kretek	Sangat Wangi	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
HARUM CENGKEH	3 PCS	SKT	Kretek	Sangat Wangi	Wangi Cengkeh	Sedang	Sedang	Berat	Plastik	Tidak Laris
DAUN DJAMBOE	20 PCS	SKT	Kretek	Sangat Wangi	Pahit	Besar	Sedang	Ringan	Plastik	Laris
REED LITE	12 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Berat	Kertas Grenjeng	Laris
REED BOLD	12 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Ringan	Kertas Grenjeng	Laris
CENGKEH ABADI	7 PCS	SKT	Kretek	Sangat Wangi	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
SOSROBAHU	5 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak Laris
REED LITE	16 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Berat	Kertas Grenjeng	Laris
SOSROBAHU	5 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak Laris
SOSROBAHU	4 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak Laris
DAUN DJAMBOE	8 PCS	SKT	Kretek	Sangat Wangi	Pahit	Besar	Sedang	Ringan	Plastik	Laris
REED BOLD	7 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Ringan	Kertas Grenjeng	Laris
SIMUSTIKA GOLD	6 PCS	SKT	Kretek	Lembut	Pedes	Sedang	Sedang	Ringan	Plastik	Laris
TUJUJ DAUN REFILL	10 PCS	SKT	Kretek	Lembut	Manis	Besar	Keras	Ringan	Kertas Grenjeng	Laris
DAUN DJAMBOE	8 PCS	SKT	Kretek	Sangat Wangi	Pahit	Besar	Sedang	Ringan	Plastik	Laris
HARUM CENGKEH	1 PCS	SKT	Kretek	Sangat Wangi	Wangi Cengkeh	Sedang	Sedang	Berat	Plastik	Tidak Laris
CENGKEH ABADI	8 PCS	SKT	Kretek	Sangat Wangi	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
REED BOLD	17 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Ringan	Kertas Grenjeng	Laris
TUJUJ DAUN REFILL	8 PCS	SKT	Kretek	Lembut	Manis	Besar	Keras	Ringan	Kertas Grenjeng	Laris
REED LITE	21 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Berat	Kertas Grenjeng	Laris

Sumber: CV Jaya Berkah Mas

**Gambar III. 1.**  
Data Penjualan Rokok

### 3.2. Tahap *Pre-Processing*

#### 3.2.1. Seleksi Data

Proses seleksi atribut dilakukan untuk memilih variabel yang akan digunakan dalam penelitian. Variabel yang digunakan pada penelitian yaitu Kecamatan, Kabupaten, Nama Produk, Kategori Rokok, Jenis Rokok, Aroma, Rasa, Ukuran Batang Rokok, Kepadatan Rokok, Hisapan Rokok, Kemasan Rokok, Harga Satuan, dan Total. Berikut pada gambar dibawah ini dataset yang sudah di seleksi.

Unit	Kategori Rokok	Jenis Rokok	Aroma	Rasa	Ukuran Batang Rokok	Kepadatan Rokok	Hisapan Rokok	Kemasan Rokok	Label
6 PCS	SKT	Kretek	Sangat Wangi	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
10 PCS	SKT	Kretek	Sangat Wangi	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
3 PCS	SKT	Kretek	Sangat Wangi	Wangi Cengkeh	Sedang	Sedang	Berat	Plastik	Tidak Laris
20 PCS	SKT	Kretek	Sangat Wangi	Pahit	Besar	Sedang	Ringan	Plastik	Laris
12 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Berat	Kertas Grenjeng	Laris
12 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Berat	Kertas Grenjeng	Laris
7 PCS	SKT	Kretek	Sangat Wangi	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
5 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak Laris
16 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Berat	Kertas Grenjeng	Laris
5 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Berat	Plastik	Tidak Laris
4 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak Laris
8 PCS	SKT	Kretek	Sangat Wangi	Pahit	Besar	Sedang	Ringan	Plastik	Laris
7 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Ringan	Kertas Grenjeng	Laris
6 PCS	SKT	Kretek	Lembut	Pedes	Sedang	Sedang	Ringan	Plastik	Laris
10 PCS	SKT	Kretek	Lembut	Manis	Besar	Keras	Ringan	Kertas Grenjeng	Laris
8 PCS	SKT	Kretek	Sangat Wangi	Pahit	Besar	Sedang	Ringan	Plastik	Laris
1 PCS	SKT	Kretek	Sangat Wangi	Wangi Cengkeh	Sedang	Sedang	Berat	Plastik	Tidak Laris
8 PCS	SKT	Kretek	Sangat Wangi	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
17 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Ringan	Kertas Grenjeng	Laris
8 PCS	SKT	Kretek	Lembut	Manis	Besar	Keras	Ringan	Kertas Grenjeng	Laris
21 PCS	SKM	Filter	Sangat Wangi	Manis	Sedang	Sedang	Berat	Kertas Grenjeng	Laris

Sumber: Penelitian Mandiri (2024)

**Gambar III. 2.**

Dataset yang telah di seleksi

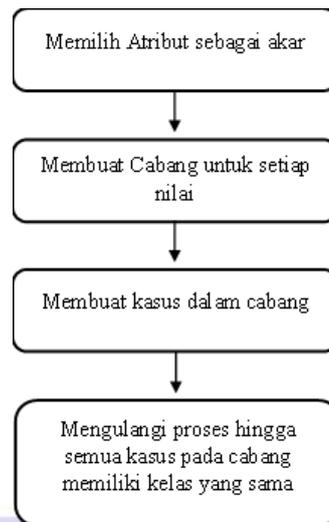
### 3.3. Pengolahan Data Menggunakan Algoritma C4.5

#### 3.3.1. *Decision Tree*

Berikut adalah langkah-langkah perhitungan dan pengolahan data menggunakan *Decision Tree*:

1. Melakukan Perhitungan menggunakan metode *Decision Tree*

Secara umum, proses membangun *Decision Tree* dilakukan melalui beberapa langkah sebagai berikut:



Sumber: Penelitian Mandiri (2024)

**Gambar III. 3.**  
Proses *Decision Tree*

- a. Pilih atribut sebagai akar.
- b. Buat cabang untuk tiap-tiap nilai.
- c. Buat Kasus dalam cabang.
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk menghitung *gain* digunakan persamaan sebagai berikut :

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

Keterangan :

S = Himpunan kasus

A = Jumlah Partisi A

|S<sub>i</sub>| = Proporsi S<sub>i</sub> terhadap S

|S| = Jumlah kasus dalam S

Sedangkan untuk menghitung nilai *entropy* digunakan persamaan sebagai berikut :

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

Keterangan

S = Himpunan kasus

n = Jumlah partisi A

$p_i$  = proporsi  $S_i$  terhadap S

1) Perhitungan *entropy* total

Tersedia total 212 kasus dalam dataset, dengan 173 kasus dalam kategori “Laris” dan 39 kasus dalam kategori “Tidak Laris”. Hasilnya, nilai total *entropy* yang dihitung adalah sebagai berikut:

$$\begin{aligned} Entropy (Total) &= \left( -\frac{173}{212} \times \log_2 \left( \frac{173}{212} \right) \right) + \left( -\frac{39}{212} \times \log_2 \left( \frac{39}{212} \right) \right) \\ &= \left( -\frac{173}{212} \times \left( \frac{\ln \left( \frac{173}{212} \right)}{\ln(2)} \right) \right) + \left( -\frac{39}{212} \times \left( \frac{\ln \left( \frac{39}{212} \right)}{\ln(2)} \right) \right) \\ &= \left( -0,81603 \left( \frac{-0,20329}{0,69314} \right) \right) + \left( -0,18396 \left( \frac{-1,69302}{0,69314} \right) \right) \\ &= ((-0,81603) \times (-0,29328)) + ((-0,18396) \times (-2,44253)) \\ &= 0,239325 + 0,44932 \\ &= \mathbf{0,68865} \end{aligned}$$

2) Perhitungan nilai *entropy* dan *gain* setiap atribut antara lain:

a) Kategori

**Tabel III. 1.**  
Tabel Atribut Kategori Rokok

Atribut	Value	Jumlah kasus (S)	LARIS (S1)	TIDAK LARIS (S2)
Kategori Rokok	SKM	44	40	4
	SKT	168	133	35

Sumber : Penelitian Mandiri (2024)

Nilai *entropy* pada masing-masing partisi:

$$\begin{aligned}
Entropy (SKM) &= \left(-\frac{40}{44} \times \log_2 \left(\frac{40}{44}\right)\right) + \left(-\frac{4}{44} \times \log_2 \left(\frac{4}{44}\right)\right) \\
&= \left(-\frac{40}{44} \times \left(\frac{\ln\left(\frac{40}{44}\right)}{\ln(2)}\right)\right) + \left(-\frac{4}{44} \times \left(\frac{\ln\left(\frac{4}{44}\right)}{\ln(2)}\right)\right) \\
&= \left(-0,90909 \left(\frac{-0,09531}{0,693147}\right)\right) + \left(-0,09091 \left(\frac{-2,39790}{0,693147}\right)\right) \\
&= ((-0,90909) \times (-0,13750)) + ((-0,09091) \times (-3,45947)) \\
&= 0,125004 + 0,31450 \\
&= 0,43950
\end{aligned}$$

$$\begin{aligned}
Entropy (SKT) &= \left(-\frac{133}{168} \times \log_2 \left(\frac{133}{168}\right)\right) + \left(-\frac{4}{168} \times \log_2 \left(\frac{4}{168}\right)\right) \\
&= \left(-\frac{133}{168} \times \left(\frac{\ln\left(\frac{133}{168}\right)}{\ln(2)}\right)\right) + \left(-\frac{4}{168} \times \left(\frac{\ln\left(\frac{4}{168}\right)}{\ln(2)}\right)\right) \\
&= \left(-0,79167 \left(\frac{-0,233614}{0,693147}\right)\right) + \left(-0,20833 \left(\frac{-1,56862}{0,693147}\right)\right) \\
&= ((-0,79167) \times (-0,23361)) + ((-0,20833) \times (-2,26305)) \\
&= 0,26682 + 0,47147 \\
&= 0,73829
\end{aligned}$$

$$\begin{aligned}
Gain \quad (Total, \quad Kategori \quad Rokok) &= \quad entropy \quad (Total) \\
&- \sum_{i=1}^n \frac{[PROSES]}{[Total]} \times Entropy (Kategori Rokok) \\
&= 0,68865 - \left(\left(\frac{44}{212} \times 0,43949\right) + \left(\frac{168}{212} \times 0,73829\right)\right) \\
&= 0,68865 - (0,09121 + 0,58505) \\
&= \mathbf{0,0124}
\end{aligned}$$

## b) Jenis Rokok

**Tabel III. 2.**  
Tabel Atribut Jenis Rokok

Atribut	Value	Jumlah kasus (S)	LARIS (S1)	TIDAK LARIS (S2)
Jenis Rokok	Filter	44	40	4
	Kretek	168	133	35

Sumber : Penelitian Mandiri (2024)

Nilai *entropy* pada masing-masing partisi:

$$\begin{aligned}
 Entropy(\text{Filter}) &= \left( -\frac{40}{44} \times \log_2 \left( \frac{40}{44} \right) \right) + \left( -\frac{4}{44} \times \log_2 \left( \frac{4}{44} \right) \right) \\
 &= \left( -\frac{40}{44} \times \left( \frac{\ln \left( \frac{40}{44} \right)}{\ln(2)} \right) \right) + \left( -\frac{4}{44} \times \left( \frac{\ln \left( \frac{4}{44} \right)}{\ln(2)} \right) \right) \\
 &= \left( -0,90909 \left( \frac{-0,09531}{0,693147} \right) \right) + \left( -0,09091 \left( \frac{-2,39790}{0,693147} \right) \right) \\
 &= \left( (-0,90909) \times (-0,13750) \right) + \left( (-0,09091) \times (-3,45947) \right) \\
 &= 0,125004 + 0,31450 \\
 &= 0,43950
 \end{aligned}$$

$$\begin{aligned}
 Entropy(\text{Kretek}) &= \left( -\frac{133}{168} \times \log_2 \left( \frac{133}{168} \right) \right) + \left( -\frac{4}{168} \times \log_2 \left( \frac{4}{168} \right) \right) \\
 &= \left( -\frac{133}{168} \times \left( \frac{\ln \left( \frac{133}{168} \right)}{\ln(2)} \right) \right) + \left( -\frac{4}{168} \times \left( \frac{\ln \left( \frac{4}{168} \right)}{\ln(2)} \right) \right) \\
 &= \left( -0,79167 \left( \frac{-0,233614}{0,693147} \right) \right) + \left( -0,20833 \left( \frac{-1,56862}{0,693147} \right) \right) \\
 &= \left( (-0,79167) \times (-0,23361) \right) + \left( (-0,20833) \times (-2,26305) \right) \\
 &= 0,26682 + 0,47147 \\
 &= 0,73829
 \end{aligned}$$

*Gain* (Total, Jenis Rokok) = *entropy* (Total)

$$- \sum_{i=1}^n \frac{[PROSES]}{[Total]} \times Entropy(\text{Kategori Rokok})$$

$$\begin{aligned}
&= 0,68865 - \left( \left( \frac{44}{212} \times 0,43949 \right) + \left( \frac{108}{212} \times 0,73829 \right) \right) \\
&= 0,68865 - (0,09121 + 0,58505) \\
&= \mathbf{0,0124}
\end{aligned}$$

c) Aroma

**Tabel III. 3.**  
Tabel Atribut Aroma

Atribut	Value	Jumlah kasus (S)	LARIS (S1)	TIDAK LARIS (S2)
Aroma	Lembut	81	52	29
	Sangat Wangi	131	121	10

Sumber : Penelitian Mandiri (2024)

Nilai *entropy* pada masing-masing partisi :

$$\begin{aligned}
Entropy \text{ (Lembut)} &= \left( -\frac{52}{81} \times \log_2 \left( \frac{52}{81} \right) \right) + \left( -\frac{29}{81} \times \log_2 \left( \frac{29}{81} \right) \right) \\
&= \left( -\frac{52}{81} \times \left( \frac{\ln \left( \frac{52}{81} \right)}{\ln(2)} \right) \right) + \left( -\frac{29}{81} \times \left( \frac{\ln \left( \frac{29}{81} \right)}{\ln(2)} \right) \right) \\
&= \left( -0,64197 \left( \frac{-0,44320}{0,693147} \right) \right) + \left( -0,35802 \left( \frac{-1,02715}{0,693147} \right) \right) \\
&= ((-0,64197) \times (-0,63941)) + ((-0,09091) \times (-1,48187)) \\
&= 0,41048 + 0,53054 \\
&= 0,94103
\end{aligned}$$

$$\begin{aligned}
Entropy \text{ (Sangat Wangi)} &= \left( -\frac{121}{131} \times \log_2 \left( \frac{121}{131} \right) \right) + \left( -\frac{10}{131} \times \log_2 \left( \frac{10}{131} \right) \right) \\
&= \left( -\frac{121}{131} \times \left( \frac{\ln \left( \frac{121}{131} \right)}{\ln(2)} \right) \right) + \left( -\frac{10}{131} \times \left( \frac{\ln \left( \frac{10}{131} \right)}{\ln(2)} \right) \right) \\
&= \left( -0,92366 \left( \frac{-0,07940}{0,693147} \right) \right) + \left( -0,07633 \left( \frac{-2,57261}{0,693147} \right) \right) \\
&= ((-0,92366) \times (-0,11455)) + ((-0,07633) \times (-3,71153))
\end{aligned}$$

$$= 0,1058 + 0,28332$$

$$= 0,38913$$

$$\text{Gain (Total, Aroma)} = \text{entropy (Total)}$$

$$- \sum_{i=1}^n \frac{[PROSES]}{[Total]} \times \text{Entropy (Kategori Rokok)}$$

$$= 0,68865 - \left( \left( \frac{81}{212} \times 0,94103 \right) + \left( \frac{131}{212} \times 0,38913 \right) \right)$$

$$= 0,68865 - (0,359544 + 0,240456)$$

$$= \mathbf{0,08867}$$

d) Rasa

**Tabel III. 4.**  
Tabel Atribut Rasa

Atribut	Value	Jumlah kasus (S)	LARIS (S1)	TIDAK LARIS (S2)
Rasa	Blueberry	34	34	0
	Manis	87	71	16
	Pahit	25	25	0
	Pedes	21	21	0
	Teh	17	0	17
	Wangi Cengkeh	28	22	6

Sumber : Penelitian Mandiri (2024)

Nilai *entropy* pada masing-masing partisi :

$$\text{Entropy (Blueberry)} = \left( - \frac{34}{34} \times \log_2 \left( \frac{34}{34} \right) \right) + \left( - \frac{0}{34} \times \log_2 \left( \frac{0}{34} \right) \right)$$

$$= \left( - \frac{34}{34} \times \left( \frac{\ln \left( \frac{34}{34} \right)}{\ln (2)} \right) \right) + \left( - \frac{0}{34} \times \left( \frac{\ln \left( \frac{0}{34} \right)}{\ln (2)} \right) \right)$$

$$= \left( -1 \left( \frac{0}{0,693147} \right) \right) + \left( 0 \left( \frac{0}{0,693147} \right) \right)$$

$$= ((-1) \times (0)) + ((0) \times (0))$$

$$= 0 + 0$$

$$= 0$$

$$\text{Entropy (Manis)} = \left( -\frac{71}{87} \times \log_2 \left( \frac{71}{87} \right) \right) + \left( -\frac{16}{87} \times \log_2 \left( \frac{16}{87} \right) \right)$$

$$= \left( -\frac{71}{87} \times \left( \frac{\ln \left( \frac{71}{87} \right)}{\ln(2)} \right) \right) + \left( -\frac{16}{87} \times \left( \frac{\ln \left( \frac{16}{87} \right)}{\ln(2)} \right) \right)$$

$$= \left( -0,81609 \left( \frac{-0,203228}{0,693147} \right) \right) + \left( -0,18390 \left( \frac{-1,6933}{0,693147} \right) \right)$$

$$= ((-0,81609) \times (-0,29319)) + ((-0,18390) \times (-2,44295))$$

$$= 0,23927 + 0,44927$$

$$= 0,68855$$

$$\text{Entropy (Pahit)} = \left( -\frac{25}{25} \times \log_2 \left( \frac{25}{25} \right) \right) + \left( -\frac{0}{25} \times \log_2 \left( \frac{0}{25} \right) \right)$$

$$= \left( -\frac{25}{25} \times \left( \frac{\ln \left( \frac{25}{25} \right)}{\ln(2)} \right) \right) + \left( -\frac{0}{34} \times \left( \frac{\ln \left( \frac{0}{25} \right)}{\ln(2)} \right) \right)$$

$$= \left( -1 \left( \frac{0}{0,693147} \right) \right) + \left( 0 \left( \frac{0}{0,693147} \right) \right)$$

$$= ((-1) \times (0)) + ((0) \times (0))$$

$$= 0 + 0$$

$$= 0$$

$$\text{Entropy (Pedes)} = \left( -\frac{21}{21} \times \log_2 \left( \frac{21}{21} \right) \right) + \left( -\frac{0}{25} \times \log_2 \left( \frac{0}{21} \right) \right)$$

$$= \left( -\frac{21}{21} \times \left( \frac{\ln \left( \frac{21}{21} \right)}{\ln(2)} \right) \right) + \left( -\frac{0}{34} \times \left( \frac{\ln \left( \frac{0}{21} \right)}{\ln(2)} \right) \right)$$

$$= \left( -1 \left( \frac{0}{0,693147} \right) \right) + \left( 0 \left( \frac{0}{0,693147} \right) \right)$$

$$= ((-1) \times (0)) + ((0) \times (0))$$

$$= 0 + 0$$

$$= 0$$

$$\begin{aligned}
Entropy (\text{Teh}) &= \left(-\frac{0}{17} \times \log_2 \left(\frac{0}{17}\right)\right) + \left(-\frac{17}{17} \times \log_2 \left(\frac{17}{17}\right)\right) \\
&= \left(-\frac{0}{17} \times \left(\frac{\ln\left(\frac{0}{17}\right)}{\ln(2)}\right)\right) + \left(-\frac{17}{17} \times \left(\frac{\ln\left(\frac{17}{17}\right)}{\ln(2)}\right)\right) \\
&= \left(0 \left(\frac{0}{0,693147}\right)\right) + \left(-1 \left(\frac{0}{0,693147}\right)\right) \\
&= ((0) \times (0)) + ((-1) \times (0)) \\
&= 0 + 0 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
Entropy (\text{Wangi Cengkeh}) &= \left(-\frac{22}{28} \times \log_2 \left(\frac{22}{28}\right)\right) + \left(-\frac{6}{28} \times \log_2 \left(\frac{6}{28}\right)\right) \\
&= \left(-\frac{22}{28} \times \left(\frac{\ln\left(\frac{22}{28}\right)}{\ln(2)}\right)\right) + \left(-\frac{6}{28} \times \left(\frac{\ln\left(\frac{6}{28}\right)}{\ln(2)}\right)\right) \\
&= \left(-0,78571 \left(\frac{-0,24116}{0,693147}\right)\right) + \left(-0,21428 \left(\frac{-1,54044}{0,693147}\right)\right) \\
&= ((-0,78571) \times (-0,34792)) + ((-0,21428) \times (-2,22240)) \\
&= 0,27337 + 0,47623 \\
&= 0,74960
\end{aligned}$$

$$\begin{aligned}
Gain \quad (\text{Total, Rasa}) &= entropy \quad (\text{Total}) \\
&- \sum_{i=1}^n \frac{[PROSES]}{[Total]} \times Entropy (\text{Kategori Rokok}) \\
&= 0,68865 - \left(\left(\frac{34}{212} \times 0\right) + \left(\frac{87}{212} \times 0,68855\right) + \left(\frac{25}{212} \times 0\right) + \left(\frac{21}{212} \times 0\right) + \right. \\
&\quad \left. \left(\frac{17}{212} \times 0\right) + \left(\frac{28}{212} \times 0,74959\right)\right) \\
&= 0,68865 - (0 + 0,24046 + 0 + 0 + 0 + 0,09900) \\
&= 0,3071
\end{aligned}$$

e) Ukuran batang rokok

**Tabel III. 5.**  
Tabel Atribut Ukuran Batang Rokok

Atribut	Value	Jumlah kasus (S)	LARIS (S1)	TIDAK LARIS (S2)
Ukuran Batang Rokok	Besar	56	56	0
	Sedang	156	117	39

Sumber : Penelitian Mandiri (2024)

Nilai *entropy* pada masing-masing partisi :

$$Entropy (\text{Bes ar}) = \left( -\frac{56}{56} \times \log_2 \left( \frac{56}{56} \right) \right) + \left( -\frac{0}{56} \times \log_2 \left( \frac{0}{56} \right) \right)$$

$$= \left( -\frac{56}{56} \times \left( \frac{\ln \left( \frac{56}{56} \right)}{\ln (2)} \right) \right) + \left( -\frac{0}{56} \times \left( \frac{\ln \left( \frac{0}{56} \right)}{\ln (2)} \right) \right)$$

$$= \left( -1 \left( \frac{0}{0,693147} \right) \right) + \left( 0 \left( \frac{0}{0,693147} \right) \right)$$

$$= ((-1) \times (0)) + ((0) \times (0))$$

$$= 0 + 0$$

$$= 0$$

$$Entropy (\text{Sed ang}) = \left( -\frac{117}{156} \times \log_2 \left( \frac{117}{156} \right) \right) + \left( -\frac{39}{156} \times \log_2 \left( \frac{39}{156} \right) \right)$$

$$= \left( -\frac{117}{156} \times \left( \frac{\ln \left( \frac{117}{156} \right)}{\ln (2)} \right) \right) + \left( -\frac{39}{156} \times \left( \frac{\ln \left( \frac{39}{156} \right)}{\ln (2)} \right) \right)$$

$$= \left( -0,75 \left( \frac{-0,28768}{0,693147} \right) \right) + \left( -0,25 \left( \frac{-1,38629}{0,693147} \right) \right)$$

$$= (-0,75 \times (-0,41503)) + ((-0,25) \times (-2,00001))$$

$$= 0,31127 + 0,50000$$

$$= 0,81128$$

*Gain* (Total, Ukuran Batang Rokok) = *entropy* (Total)

$$- \sum_{i=1}^n \frac{[PROSES]}{[Total]} \times Entropy (\text{Kategori Rokok})$$

$$\begin{aligned}
&= 0,68865 - \left( \left( \frac{56}{212} \times 0 \right) + \left( \frac{156}{212} \times 0,81127 \right) \right) \\
&= 0,68865 - (0 + 0,59698) \\
&= \mathbf{0,0917}
\end{aligned}$$

f) Kepadatan rokok

**Tabel III. 6.**  
Tabel Atribut Kepadatan Rokok

Atribut	Value	Jumlah kasus (S)	LARIS (S1)	TIDAK LARIS (S2)
Kepadatan Rokok	Keras	104	87	17
	Sedang	108	86	22

Sumber : Penelitian Mandiri (2024)

Nilai *entropy* pada masing-masing partisi :

$$\begin{aligned}
Entropy \text{ (Keras)} &= \left( -\frac{87}{104} \times \log_2 \left( \frac{87}{104} \right) \right) + \left( -\frac{17}{104} \times \log_2 \left( \frac{17}{104} \right) \right) \\
&= \left( -\frac{87}{104} \times \left( \frac{\ln \left( \frac{87}{104} \right)}{\ln(2)} \right) \right) + \left( -\frac{17}{104} \times \left( \frac{\ln \left( \frac{17}{104} \right)}{\ln(2)} \right) \right) \\
&= \left( -0,83653 \left( \frac{-0,17848}{0,693147} \right) \right) + \left( -0,16346 \left( \frac{-1,81117}{0,693147} \right) \right) \\
&= ((-0,83653) \times (-0,25749)) + ((-0,16346) \times (-2,61299)) \\
&= 0,21540 + 0,42712 \\
&= \mathbf{0,64252}
\end{aligned}$$

$$\begin{aligned}
Entropy \text{ (Sedang)} &= \left( -\frac{86}{108} \times \log_2 \left( \frac{86}{108} \right) \right) + \left( -\frac{22}{108} \times \log_2 \left( \frac{22}{108} \right) \right) \\
&= \left( -\frac{86}{108} \times \left( \frac{\ln \left( \frac{86}{108} \right)}{\ln(2)} \right) \right) + \left( -\frac{22}{108} \times \left( \frac{\ln \left( \frac{22}{108} \right)}{\ln(2)} \right) \right) \\
&= \left( -0,79629 \left( \frac{-0,22778}{0,693147} \right) \right) + \left( -0,20370 \left( \frac{-1,59108}{0,693147} \right) \right) \\
&= ((-0,79629) \times (-0,32862)) + ((-0,20833) \times (-2,29546)) \\
&= \mathbf{0,26167 + 0,46759}
\end{aligned}$$

$$= 0,72927$$

*Gain* (Total, Kepadatan Rokok) = *entropy* (Total)

$$- \sum_{i=1}^n \frac{[PROSES]}{[Total]} \times \text{Entropy (Kategori Rokok)}$$

$$= 0,68865 - \left( \left( \frac{104}{212} \times 0,64252 \right) + \left( \frac{108}{212} \times 0,729273 \right) \right)$$

$$= 0,68865 - (0,31520 + 0,37151)$$

$$= \mathbf{0,00195}$$

g) Hisapan rokok

**Tabel III. 7.**  
Tabel Atribut Hisapan Rokok

Atribut	Value	Jumlah kasus (S)	LARIS (S1)	TIDAK LARIS (S2)
Hisapan Rokok	Berat	17	7	10
	Ringan	195	166	29

Sumber : Penelitian Mandiri (2024)

Nilai *entropy* pada masing-masing partisi :

$$\text{Entropy (Berat)} = \left( -\frac{7}{17} \times \log_2 \left( \frac{7}{17} \right) \right) + \left( -\frac{10}{17} \times \log_2 \left( \frac{10}{17} \right) \right)$$

$$= \left( -\frac{7}{17} \times \left( \frac{\ln \left( \frac{7}{17} \right)}{\ln (2)} \right) \right) + \left( -\frac{10}{17} \times \left( \frac{\ln \left( \frac{10}{17} \right)}{\ln (2)} \right) \right)$$

$$= \left( -0,41176 \left( \frac{-0,8873}{0,693147} \right) \right) + \left( -0,58823 \left( \frac{-0,53062}{0,693147} \right) \right)$$

$$= ((-0,41176) \times (-1,28011)) + ((-0,58823) \times (-0,76554))$$

$$= 0,52710 + 0,45032$$

$$= 0,977425$$

$$\text{Entropy (Ringan)} = \left( -\frac{166}{195} \times \log_2 \left( \frac{166}{195} \right) \right) + \left( -\frac{29}{195} \times \log_2 \left( \frac{29}{195} \right) \right)$$

$$= \left( -\frac{166}{195} \times \left( \frac{\ln \left( \frac{166}{195} \right)}{\ln (2)} \right) \right) + \left( -\frac{29}{195} \times \left( \frac{\ln \left( \frac{29}{195} \right)}{\ln (2)} \right) \right)$$

$$\begin{aligned}
&= \left( -0,85128 \left( \frac{-0,16101}{0,693147} \right) \right) + \left( -0,14871 \left( \frac{-1,90570}{0,693147} \right) \right) \\
&= \left( (-0,85128) \times (-0,23229) \right) + \left( (-0,14871) \times (-2,74937) \right) \\
&= 0,197746 + 0,408881 \\
&= 0,60663
\end{aligned}$$

*Gain* (Total, Hisapan Rokok) = *entropy* (Total)

$$- \sum_{i=1}^n \frac{[PROSES]}{[Total]} \times Entropy (Kategori Rokok)$$

$$\begin{aligned}
&= 0,68865 - \left( \left( \frac{17}{212} \times 0,97741 \right) + \left( \frac{195}{212} \times 0,60662 \right) \right) \\
&= 0,68865 - (0,07838 + 0,55798) \\
&= \mathbf{0,05231}
\end{aligned}$$

h) Kemasan rokok

**Tabel III. 8.**  
Tabel Atribut Kemasan Rokok

Atribut	Value	Jumlah kasus (S)	LARIS (S1)	TIDAK LARIS (S2)
Kemasan Rokok	Kertas Grenjeng	75	71	4
	Plastik	137	102	35

Sumber : Penelitian Mandiri (2024)

Nilai *entropy* pada masing-masing partisi :

$$\begin{aligned}
Entropy (\text{Kertas Grenjeng}) &= \left( - \frac{71}{75} \times \log_2 \left( \frac{71}{75} \right) \right) + \left( - \frac{4}{75} \times \log_2 \left( \frac{4}{75} \right) \right) \\
&= \left( - \frac{71}{75} \times \left( \frac{\ln \left( \frac{71}{75} \right)}{\ln (2)} \right) \right) + \left( - \frac{4}{75} \times \left( \frac{\ln \left( \frac{4}{75} \right)}{\ln (2)} \right) \right) \\
&= \left( -0,94666 \left( \frac{-0,054808}{0,693147} \right) \right) + \left( -0,05333 \left( \frac{-2,931193}{0,693147} \right) \right) \\
&= \left( (-0,94666) \times (-0,07907) \right) + \left( (-0,05333) \times (-4,22886) \right)
\end{aligned}$$

$$= 0,074855 + 0,22553$$

$$= 0,300394$$

$$Entropy (\text{Plastik}) = \left( -\frac{102}{137} \times \log_2 \left( \frac{102}{137} \right) \right) + \left( -\frac{35}{137} \times \log_2 \left( \frac{35}{137} \right) \right)$$

$$= \left( -\frac{102}{137} \times \left( \frac{\ln \left( \frac{102}{137} \right)}{\ln(2)} \right) \right) + \left( -\frac{35}{137} \times \left( \frac{\ln \left( \frac{35}{137} \right)}{\ln(2)} \right) \right)$$

$$= \left( -0,74452 \left( \frac{-0,295008}{0,693147} \right) \right) + \left( -0,25547 \left( \frac{-1,36463}{0,693147} \right) \right)$$

$$= \left( (-0,74452) \times (-0,42561) \right) + \left( (-0,25547) \times (-1,96876) \right)$$

$$= 0,31687 + 0,50296$$

$$= 0,81984$$

$$Gain (\text{Total, Kemasan Rokok}) = entropy (\text{Total})$$

$$- \sum_{i=1}^n \frac{[PROSES]}{[Total]} \times Entropy (\text{Kategori Rokok})$$

$$= 0,68865 - \left( \left( \frac{75}{212} \times 0,30039 \right) + \left( \frac{137}{212} \times 0,81984 \right) \right)$$

$$= 0,68865 - (0,10627 + 0,52980)$$

$$= \mathbf{0,052595}$$

3) Penentuan *Node* ke 1 pohon keputusan

Berikut adalah hasil perhitungan nilai gain :

**Tabel III. 9.**

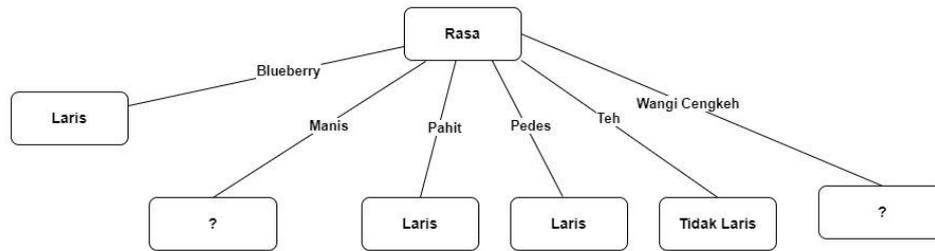
Tabel Penentuan *Node* 1 Pohon Keputusan

<i>Node</i>	Atribut	Jml Kasus (S)	Laris (Si)	Tidak Laris (Si)	Entropy	Gain	
		212	173	39	<b>0,689</b>		
1	<b>Kategori Rokok</b>	SKM	44	40	4	<b>0,439</b>	0,0124
		SKT	168	133	35	<b>0,738</b>	
		Jenis Rokok					
2	Filter	44	40	4	<b>0,439</b>	0,0124	

		Kretek	168	133	35	<b>0,738</b>	
4	Aroma						0,0887
		Lembut	81	52	29	<b>0,941</b>	
		Tidak	131	121	10	<b>0,389</b>	
5	Rasa						<b>0,3071</b>
		Blueberry	34	34	0	<b>0</b>	
		Manis	87	71	16	<b>0,688552168</b>	
		Pahit	25	25	0	<b>0</b>	
		Pedes	21	21	0	<b>0</b>	
		Teh	17	0	17	<b>0</b>	
		Wangi Cengkeh	28	22	6	<b>0,750</b>	
6	Ukuran Batang Rokok						0,0917
		Besar	56	56	0	<b>0,000</b>	
		Sedang	156	117	39	<b>0,811</b>	
7	Kepadatan Rokok						- 0,0345
		Keras	104	87	17	<b>0,717</b>	
		Sedang	108	86	22	<b>0,729</b>	
8	Hisapan Rokok						0,0523
		Berat	17	7	10	<b>0,977</b>	
		Ringan	195	166	29	<b>0,607</b>	
9	Kemasan Rokok						0,0526
		Kertas Grenjeng	75	71	4	<b>0,300</b>	
		Plastik	137	102	35	<b>0,820</b>	

Sumber: Penelitian Mandiri (2024)

Berdasarkan hasil perhitungan nilai gain tersebut, kemudian pilih gain terbesar yang kemudian digunakan sebagai *root Node* (akar) dari pohon keputusan, diketahui nilai gain terbesar adalah Rasa, yakni sebesar 0,3071 sehingga atribut Rasa digunakan Sebagai *root*. Maka, pohon keputusan dapat dibentuk seperti berikut :



Sumber: Penelitian Mandiri (2024)

**Gambar III. 4.**  
Atribut Rasa sebagai *root* pohon keputusan

4) Penentuan *Node* ke 2 pohon keputusan

Berikut adalah hasil perhitungan nilai *gain* :

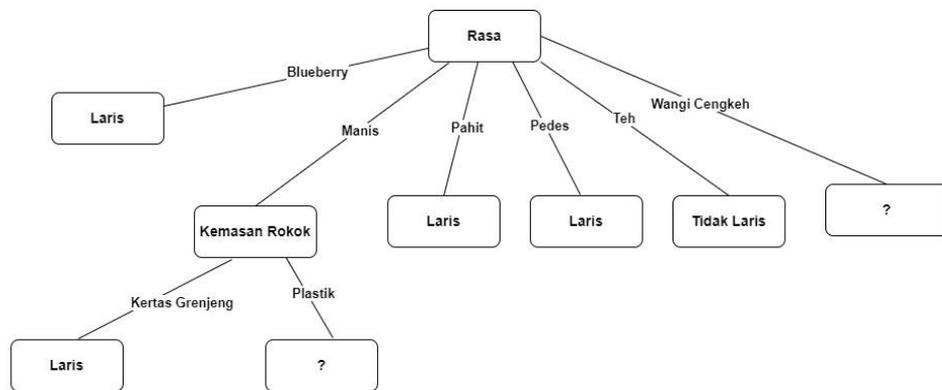
**Tabel III. 10.**  
Tabel Penentuan *Node* 2 Pohon Keputusan

<i>Node</i>	Atribut	Jml Kasus (S)	Laris (Si)	Tidak Laris (Si)	Entropy	Gain
	Rasa Manis	87	71	16	<b>0,689</b>	
1	<b>Kategori Rokok</b>					0,0441
	SKM	44	40	4	<b>0,439</b>	
	SKT	43	31	12	<b>0,854</b>	
2	Jenis Rokok					0,0441
	Filter	44	40	4	<b>0,439</b>	
	Kretek	43	31	12	<b>0,854</b>	
4	Aroma					0,0480
	Lembut	42	30	12	<b>0,863</b>	
	Sangat Wangi	45	41	4	<b>0,433</b>	
6	Ukuran Batang Rokok					0,1275
	Besar	30	30	0	<b>0,000</b>	
	Sedang	57	41	16	<b>0,856</b>	
7	Kepadatan Rokok					0,1625
	Keras	36	36	0	<b>0,000</b>	
	Sedang	51	35	16	<b>0,897</b>	
8	Hisapan Rokok					0,0193
	Berat	11	7	4	<b>0,946</b>	
	Ringan	76	64	12	<b>0,629</b>	

9	Kemasan Rokok					0,2005	
		Kertas Grenjeng	66	62	4		0,330
		Plastik	21	9	12		0,985

Sumber: Penelitian Mandiri (2024)

Berdasarkan hasil perhitungan nilai gain tersebut, kemudian pilih gain terbesar yang kemudian digunakan sebagai *root Node* (akar) dari pohon keputusan, diketahui nilai gain terbesar adalah Kemasan Rokok, yakni sebesar 0,2005 sehingga atribut Kemasan Rokok akan digunakan Sebagai *root*. Maka, pohon keputusan dapat dibentuk seperti berikut :



Sumber: Penelitian Mandiri (2024)

**Gambar III. 5.**  
Atribut Kemasan Rokok sebagai *root* pohon keputusan

5) Penentuan *Node* ke 3 pohon keputusan

Berikut adalah hasil perhitungan nilai gain :

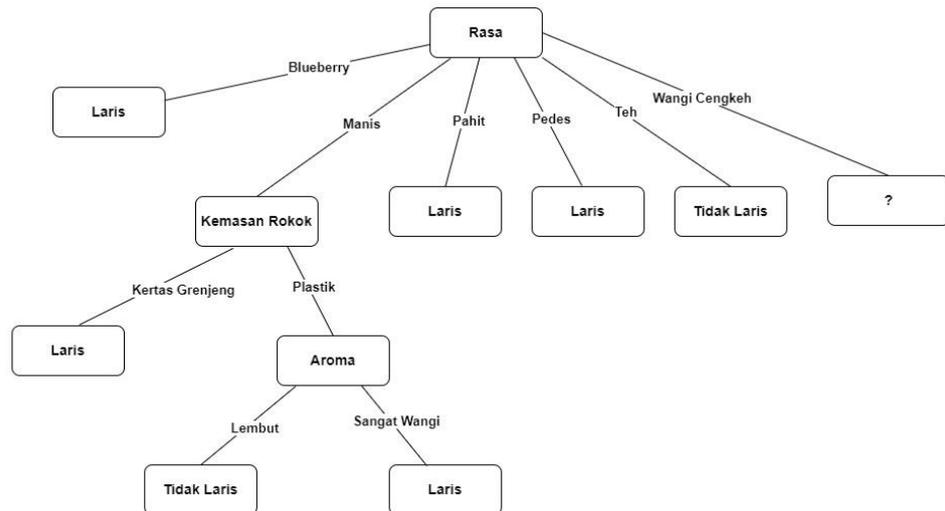
**Tabel III. 11.**  
Tabel Penentuan *Node* 3 Pohon Keputusan

<i>Node</i>	Atribut	Jml Kasus (S)	Laris (Si)	Tidak Laris (Si)	Entropy	Gain
	Rasa Manis dan Kemasan Rokok Plastik	21	9	12	0,985	
1	<b>Kategori Rokok</b>					0,4696

		SKM	6	6	0	<b>0,000</b>	
		SKT	15	3	12	<b>0,722</b>	
2	Jenis Rokok						0,469 6
		Filter	6	6	0	<b>0,000</b>	
		Kretek	15	3	12	<b>0,722</b>	
4	Aroma						<b>0,743</b> 0
		Lembut	13	1	12	<b>0,391</b>	
		Sangat Wangi	8	8	0	<b>0,000</b>	
6	Ukuran Batang Rokok						0,060 5
		Besar	1	1	0	<b>0,000</b>	
		Sedang	20	8	12	<b>0,971</b>	
7	Kepadatan Rokok						0,590 8
		Keras	7	7	0	<b>0,000</b>	
		Sedang	14	2	12	<b>0,592</b>	
8	Hisapan Rokok						0,000 0
		Berat	21	9	12	<b>0,985</b>	
		Ringan	0	0	0	<b>0,000</b>	

Sumber: Penelitian Mandiri (2024)

Berdasarkan hasil perhitungan nilai gain tersebut, kemudian pilih gain terbesar yang kemudian digunakan sebagai *root Node* (akar) dari pohon



Sumber: Penelitian Mandiri (2024)

**Gambar III. 6.**  
Atribut Aroma sebagai *root* pohon keputusan

6) Penentuan *Node* ke 4 pohon keputusan

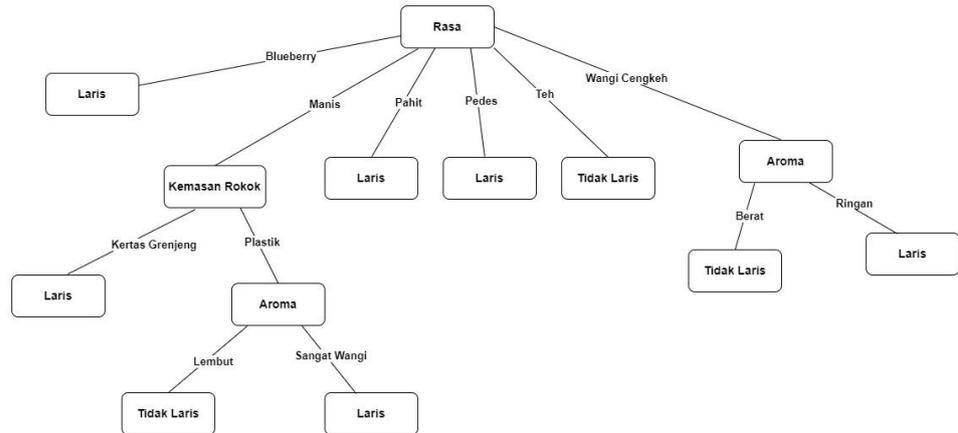
Berikut adalah hasil perhitungan nilai gain :

**Tabel III. 12.**  
Tabel Penentuan *Node* 4 Pohon Keputusan

<i>Node</i>	Atribut	Jml Kasus (S)	Laris (Si)	Tidak Laris (Si)	Entropy	Gain	
	Rasa Wangi Cengkeh	28	22	6	<b>0,750</b>		
1	<b>Kategori Rokok</b>					0,7496	
		SKM	0	0	0	<b>0,000</b>	
		SKT	28	28	0	<b>0,000</b>	
2	Jenis Rokok					0,7496	
		Filter	0	0	0	<b>0,000</b>	
		Kretek	28	28	0	<b>0,000</b>	
4	Aroma					0,0127	
		Lembut	1	0	0	<b>0,000</b>	
		Sangat Wangi	27	21	6	<b>0,764</b>	
6	Ukuran Batang Rokok					0,0397	
		Besar	3	3	0	<b>0,000</b>	
		Sedang	25	19	6	<b>0,795</b>	
7	Kepadatan Rokok					0,2294	
		Keras	13	13	0	<b>0,000</b>	
		Sedang	15	9	6	<b>0,971</b>	
8	Hisapan Rokok					<b>0,7496</b>	
		Berat	6	0	6	<b>0,000</b>	
		Ringan	22	22	0	<b>0,000</b>	
9	Kemasan Rokok					0,0694	
		Kertas Grenjeng	5	5	0	<b>0,000</b>	
		Plastik	23	17	6	<b>0,828</b>	

Sumber: Penelitian Mandiri (2024)

Berdasarkan hasil perhitungan nilai gain tersebut, kemudian pilih gain terbesar yang kemudian digunakan sebagai *root Node* (akar) dari pohon keputusan, diketahui nilai gain terbesar adalah Hisapan Rokok, yakni sebesar 0,7496 sehingga atribut Hisapan Rokok akan digunakan Sebagai *root*. Maka, pohon keputusan dapat dibentuk seperti berikut :



Sumber: Penelitian Mandiri (2024)

**Gambar III. 7.**  
Atribut Hisapan Rokok sebagai *root* pohon keputusan

2. Pengolahan Data Dengan *Decision Tree* Menggunakan Rapidminer

Berikut adalah model *Decision Tree* yang telah dibuat dengan menganalisis data menggunakan aplikasi Rapidminer:

- a. Pertama memasukan dataset untuk pengolahan *Decision Tree*.

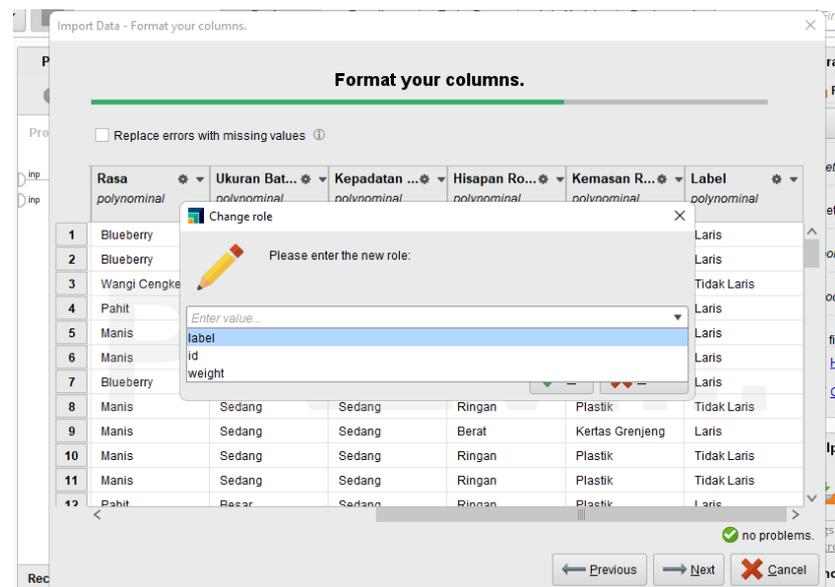
	A	B	C	D	E	F	G	H	I	J
1	Unit	Kategori...	Jenis Ro...	Aroma	Rasa	Ukuran ...	Kepadat...	Hisapan ...	Kemasa...	Label
2	6 PCS	SKT	Kretek	Sangat ...	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
3	10 PCS	SKT	Kretek	Sangat ...	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
4	3 PCS	SKT	Kretek	Sangat ...	Wangi C...	Sedang	Sedang	Berat	Plastik	Tidak L
5	20 PCS	SKT	Kretek	Sangat ...	Pahit	Besar	Sedang	Ringan	Plastik	Laris
6	12 PCS	SKM	Filter	Sangat ...	Manis	Sedang	Sedang	Berat	Kertas G...	Laris
7	12 PCS	SKM	Filter	Sangat ...	Manis	Sedang	Sedang	Ringan	Kertas G...	Laris
8	7 PCS	SKT	Kretek	Sangat ...	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
9	5 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak L
10	16 PCS	SKM	Filter	Sangat ...	Manis	Sedang	Sedang	Berat	Kertas G...	Laris
11	5 PCS	SKT	Kretek	Sangat ...	Manis	Sedang	Sedang	Ringan	Plastik	Tidak L
12	4 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak L
13	8 PCS	SKT	Kretek	Sangat ...	Pahit	Besar	Sedang	Ringan	Plastik	Laris

Sumber: Penelitian Mandiri (2024)

**Gambar III. 8.**  
Dataset Penjualan Rokok

Dataset yang dimasukkan untuk pengolahan data yaitu dari dataset pengujian penjualan rokok sebanyak 212 data uji, dengan atribut unit, kategori, jenis rokok, aroma, rasa, ukuran rokok, kepadatan rokok, hisapan rokok, kemasan rokok, dan label.

- b. Kemudian mengubah *role* menjadi label yang digunakan pada aplikasi Rapidminer.

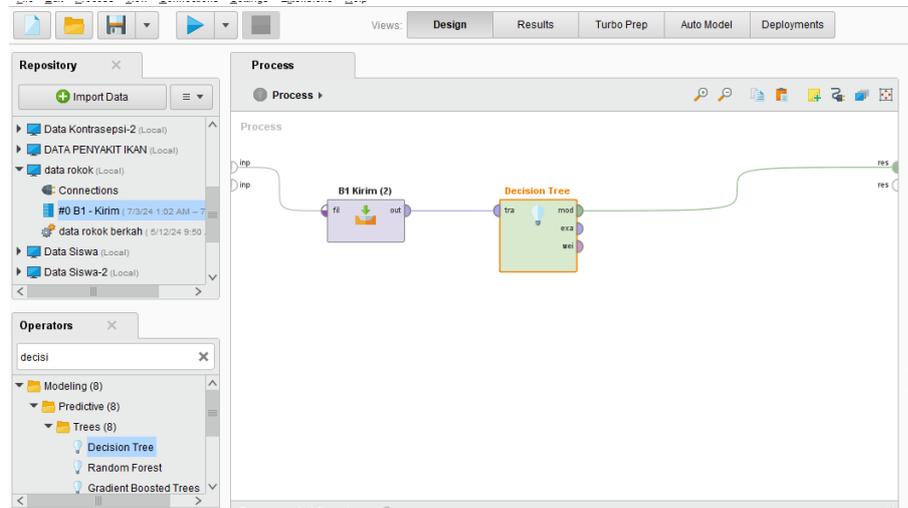


Sumber: Penelitian Mandiri (2024)

**Gambar III. 9.**  
Penambahan Label Pada Atribut Label

Pada tahap ini, atribut label diubah menjadi rule label untuk menentukan kelas dari suatu data.

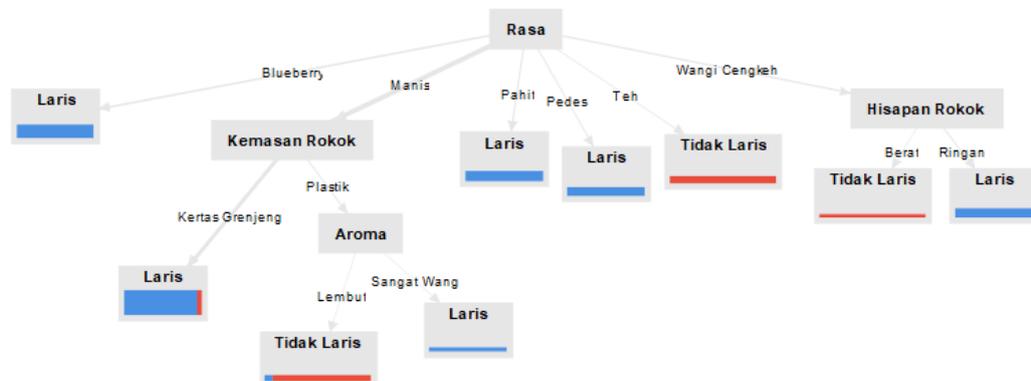
- c. Selanjutnya memasukkan *operator Decision Tree* dengan mengetik di kolom *search for operator* untuk mengolah dataset menjadi model *Decision Tree*.



Sumber: Penelitian Mandiri (2024)

**Gambar III. 10.**  
Pengolahan *Decision Tree* Menggunakan Rapidminer

- d. Berikut adalah hasil *Decision Tree* yang dibuat dengan menganalisis data menggunakan aplikasi Rapidminer.



Sumber: Penelitian Mandiri (2024)

**Gambar III. 11.** Hasil *Decision Tree*  
pada Rapidminer

Dari pohon keputusan pada gambar III.11 didapatkan *Rule* untuk klasifikasi penjualan laris. *Rule* yang sebagai berikut :

### Tree

```
Rasa = Blueberry: Laris {Laris=34, Tidak Laris=0}
Rasa = Manis
| Kemasan Rokok = Kertas Grenjeng: Laris {Laris=62, Tidak Laris=4}
| Kemasan Rokok = Plastik
| | Aroma = Lembut: Tidak Laris {Laris=1, Tidak Laris=12}
| | Aroma = Sangat Wangi: Laris {Laris=8, Tidak Laris=0}
Rasa = Pahit: Laris {Laris=25, Tidak Laris=0}
Rasa = Pedes: Laris {Laris=21, Tidak Laris=0}
Rasa = Teh: Tidak Laris {Laris=0, Tidak Laris=17}
Rasa = Wangi Cengkeh
| Hisapan Rokok = Berat: Tidak Laris {Laris=0, Tidak Laris=6}
| Hisapan Rokok = Ringan: Laris {Laris=22, Tidak Laris=0}
```

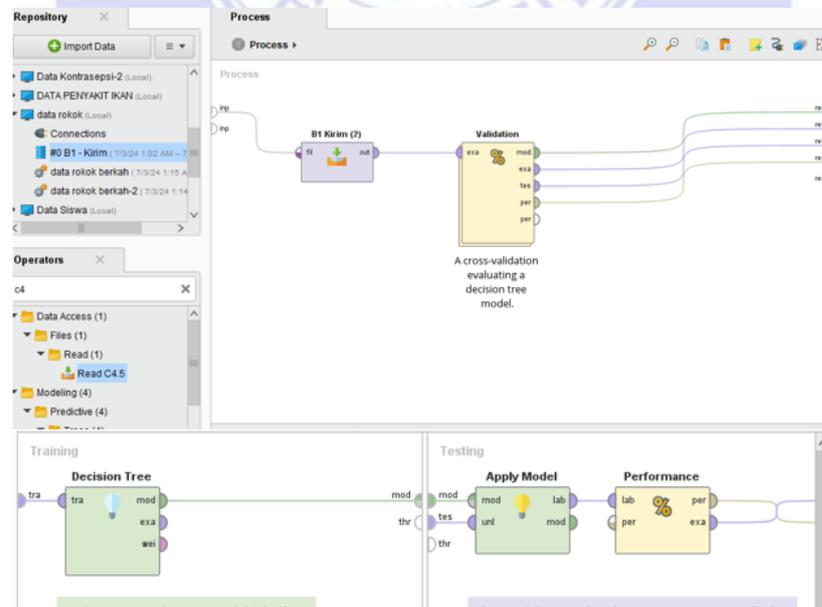
Sumber: Penelitian Mandiri (2024)

**Gambar III. 12.**  
*Rule* hasil algoritma *Decision Tree*

Deskripsi dari *Rule* pohon keputusan adalah sebagai berikut :

- 1) Jika produk rokok memiliki Rasa Blueberry, maka produk tersebut Laris. Perbandingan Laris dan Tidak Laris adalah 34:0.
- 2) Jika produk rokok memiliki Rasa Manis, dan Kategori Kemasan Rokok menggunakan Kertas Grenjeng, maka produk tersebut Laris. Perbandingan Laris dan Tidak Laris adalah 62:4.
- 3) Jika produk rokok memiliki Rasa Manis, Kemasan Rokok menggunakan Plastik dengan Kategori Aroma Lembut, maka produk tersebut Tidak Laris. Perbandingan Laris dan Tidak Laris adalah 1:12.
- 4) Jika produk rokok memiliki Rasa Manis, Kemasan Rokok menggunakan Plastik dengan Kategori Aroma Sangat Wangi, maka produk tersebut Laris. Perbandingan Laris dan Tidak Laris adalah 8:0.
- 5) Jika produk rokok memiliki Rasa Pahit, maka produk tersebut Laris. Perbandingan Laris dan Tidak Laris adalah 25:0.

- 6) Jika produk rokok memiliki Rasa Pedes, maka produk tersebut Laris. Perbandingan Laris dan Tidak Laris adalah 21:0.
  - 7) Jika produk rokok memiliki Rasa Teh, maka produk tersebut Tidak Laris. Perbandingan Laris dan Tidak Laris adalah 0:17.
  - 8) Jika produk rokok memiliki Rasa Wangi Cengkeh, dan Kategori Hisapan Rokok Berat, maka produk tersebut Tidak Laris. Perbandingan Laris dan Tidak Laris adalah 0:6.
  - 9) Jika produk rokok memiliki Rasa Wangi Cengkeh, dan Kategori Hisapan Rokok Ringan, maka produk tersebut Tidak Laris. Perbandingan Laris dan Tidak Laris adalah 22:0.
- e. Untuk mengetahui akurasi data dan menampilkan *Area Under Curve (AUC)* menggunakan aplikasi Rapidminer dengan cara menghubungkan dataset ke *validation*, di dalam *validation* terhubung antara *decision tree*, *apply model* dan *performance*.



Sumber: Penelitian Mandiri (2024)

**Gambar III. 13.**  
Proses Pengolahan Akurasi Data pada Rapidminer

f. Berikut adalah hasil akurasi data dan *Area Under Curve (AUC)*

	true Laris	true Tidak Laris	class precision
pred. Laris	172	6	96.63%
pred. Tidak Laris	1	33	97.06%
class recall	99.42%	84.62%	

accuracy: 96.73% +/- 3.11% (micro average: 96.70%)

Sumber: Penelitian Mandiri (2024)

**Gambar III. 14.**  
Nilai *Accuracy* pada *Decision Tree*

Pada **Gambar III.14** terlihat bahwa hasil nilai akurasi dari *Decision Tree* sebesar 96.73% dengan  $\pm 3.11\%$  dan *micro average* 96.70%. Performa nilai akurasi yang didapatkan sudah sangat baik.

	true Laris	true Tidak Laris	class precision
pred. Laris	172	6	96.63%
pred. Tidak Laris	1	33	97.06%
class recall	99.42%	84.62%	

precision: 98.00% +/- 6.32% (micro average: 97.06%) (positive class: Tidak Laris)

Sumber: Penelitian Mandiri (2024)

**Gambar III. 15.**  
Nilai *Precision* pada *Decision Tree*

Pada **Gambar III.15** terlihat bahwa hasil nilai presisi dari *Decision Tree* sebesar 98.00% dengan  $\pm 6.32\%$  dan *micro average* 97.06% dengan kasus kelas “Tidak Laris” jadi *positive*.

	true Laris	true Tidak Laris	class precision
pred. Laris	172	6	96.63%
pred. Tidak Laris	1	33	97.06%
class recall	99.42%	84.62%	

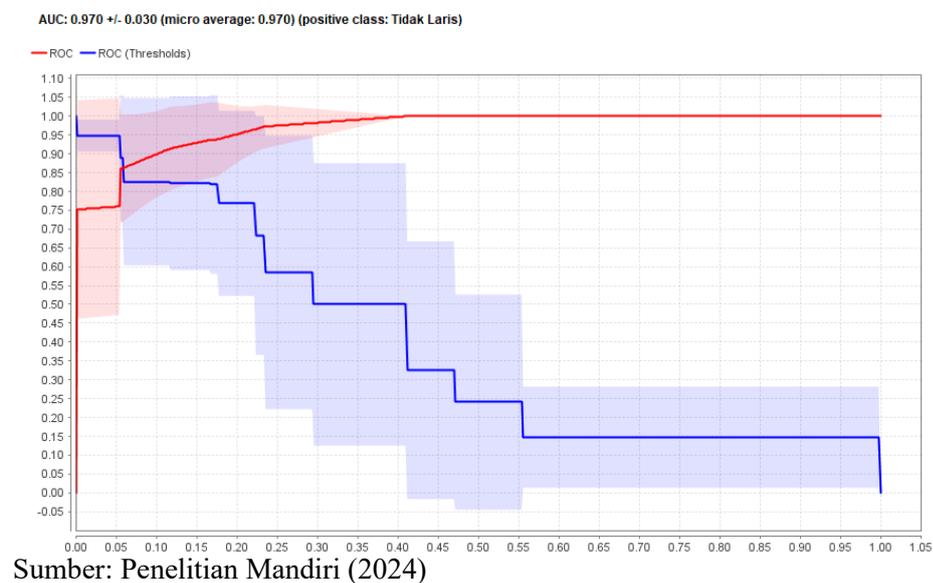
recall: 85.17% +/- 16.28% (micro average: 84.62%) (positive class: Tidak Laris)

Sumber: Penelitian Mandiri (2024)

**Gambar III. 16.**  
Nilai *Recall* pada *Decision Tree*

Pada **Gambar III.16** terlihat bahwa hasil nilai recall dari *Decision Tree* sebesar 85.17% dengan  $\pm 16.28\%$  dan *micro average* 84.62% dengan kasus kelas “Tidak Laris” jadi *positive*. Dari ketiga gambar diatas, dapat dilihat distribusi hasil prediksi dan data sebenarnya:

- 1) **True Laris** (data sebenarnya "Laris" yang diprediksi benar) sebanyak 172
- 2) **True Tidak Laris** (data sebenarnya "Tidak Laris" yang diprediksi benar) sebanyak 33
- 3) **False Laris** (data sebenarnya "Tidak Laris" yang diprediksi sebagai "Laris") sebanyak 6
- 4) **False Tidak Laris** (data sebenarnya "Laris" yang diprediksi sebagai "Tidak Laris") sebanyak 1



**Gambar III. 17.**

Nilai *Area Under Curve* (AUC) pada *Decision Tree*

Pada **Gambar III.17** terlihat bahwa hasil *Area Under Curve* dari *Decision Tree* sebesar 0.970 dengan  $\pm 0.030$ , termasuk kedalam kategori sangat baik.

### 3.3.2. Algoritma C4.5

Berikut adalah model Algoritma C4.5 yang telah dibuat dengan menganalisis data menggunakan aplikasi Rapidminer:

1. Pertama memasukkan dataset untuk pengolahan Algoritma C4.5.

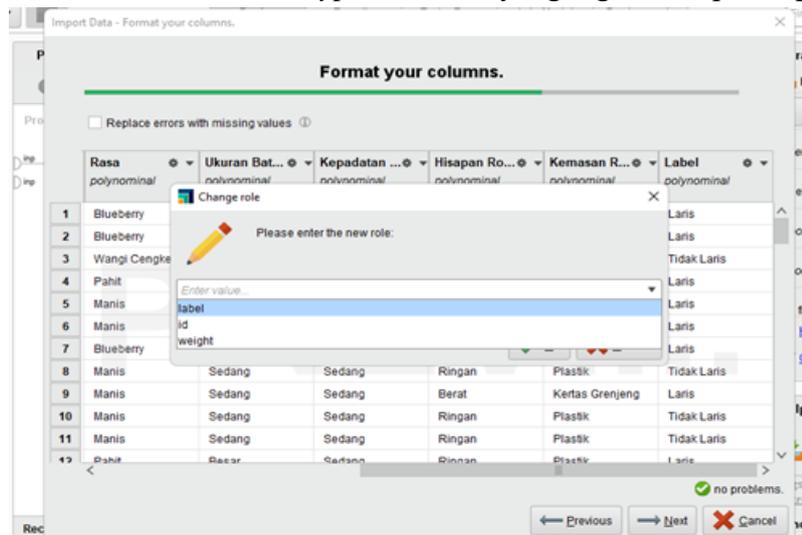
	A	B	C	D	E	F	G	H	I	J
1	Unit	Kategori	Jenis Ro...	Aroma	Rasa	Ukuran ...	Kepadat...	Hisapan ...	Kemas...	Label
2	6 PCS	SKT	Kretek	Sangat ...	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
3	10 PCS	SKT	Kretek	Sangat ...	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
4	3 PCS	SKT	Kretek	Sangat ...	Wangi C...	Sedang	Sedang	Berat	Plastik	Tidak L
5	20 PCS	SKT	Kretek	Sangat ...	Pahit	Besar	Sedang	Ringan	Plastik	Laris
6	12 PCS	SKM	Filter	Sangat ...	Manis	Sedang	Sedang	Berat	Kertas G...	Laris
7	12 PCS	SKM	Filter	Sangat ...	Manis	Sedang	Sedang	Ringan	Kertas G...	Laris
8	7 PCS	SKT	Kretek	Sangat ...	Blueberry	Sedang	Keras	Ringan	Plastik	Laris
9	5 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak L
10	16 PCS	SKM	Filter	Sangat ...	Manis	Sedang	Sedang	Berat	Kertas G...	Laris
11	5 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak L
12	4 PCS	SKT	Kretek	Lembut	Manis	Sedang	Sedang	Ringan	Plastik	Tidak L
13	8 PCS	SKT	Kretek	Sangat ...	Pahit	Besar	Sedang	Ringan	Plastik	Laris

Sumber: Penelitian Mandiri (2024)

**Gambar III. 18.**  
Dataset Penjualan Rokok

Dataset yang dimasukkan untuk pengolahan data yaitu dari dataset pengujian penjualan rokok sebanyak 212 data uji, dengan atribut unit, kategori, jenis rokok, aroma, rasa, ukuran rokok, kepadatan rokok, hisapan rokok, kemasan rokok, dan label.

2. Kemudian menentukan *type* dan label yang digunakan pada aplikasi Rapidminer.

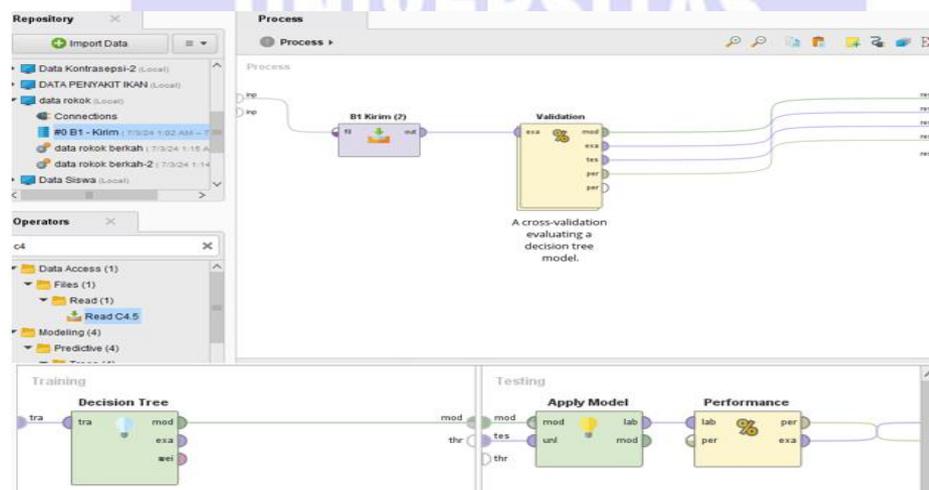


Sumber: Penelitian Mandiri (2024)

**Gambar III. 19.**  
Penambahan Label Pada Atribut Label

Pada tahap ini, atribut label diubah menjadi rule label untuk menentukan kelas dari suatu data.

3. Untuk mengetahui akurasi data dan menampilkan *Area Under Curve* (AUC) menggunakan aplikasi Rapidminer dengan cara menghubungkan dataset ke *validation*, di dalam *validation* terhubung antara *decision tree*, *apply model* dan *performance*.



Sumber: Penelitian Mandiri (2024)

**Gambar III. 20.**  
Proses Pengolahan Akurasi Data pada Rapidminer

4. Berikut adalah hasil akurasi data dan *Area Under Curve (AUC)*

Criterion	Table View	Plot View
accuracy	accuracy: 96.73% +/- 3.11% (micro average: 96.70%)	
precision		
recall		
AUC (optimistic)		
AUC		
AUC (pessimistic)		

	true Laris	true Tidak Laris	class precision
pred. Laris	172	6	96.63%
pred. Tidak Laris	1	33	97.06%
class recall	99.42%	84.62%	

Sumber: Penelitian Mandiri (2024)

**Gambar III. 21.**  
Nilai *Accuracy* pada Algoritma C4.5

Pada **Gambar III.21** terlihat bahwa hasil nilai akurasi dari *Decision Tree* sebesar 96.73% dengan  $\pm 3.11\%$  dan *micro average* 96.70%. Performa nilai akurasi yang didapatkan sudah sangat baik.

Criterion	Table View	Plot View
accuracy	precision: 98.00% +/- 6.32% (micro average: 97.06%) (positive class: Tidak Laris)	
precision		
recall		
AUC (optimistic)		
AUC		
AUC (pessimistic)		

	true Laris	true Tidak Laris	class precision
pred. Laris	172	6	96.63%
pred. Tidak Laris	1	33	97.06%
class recall	99.42%	84.62%	

Sumber: Penelitian Mandiri (2024)

**Gambar III. 22.**  
Nilai *Precision* pada Algoritma C4.5

Pada **Gambar III.22** terlihat bahwa hasil nilai presisi dari *Decision Tree* sebesar 98.00% dengan  $\pm 6.32\%$  dan *micro average* 97.06% dengan kasus kelas “Tidak Laris” jadi *positive*.

Criterion	Table View	Plot View
accuracy	recall: 85.17% +/- 16.28% (micro average: 84.62%) (positive class: Tidak Laris)	
precision		
recall		
AUC (optimistic)		
AUC		
AUC (pessimistic)		

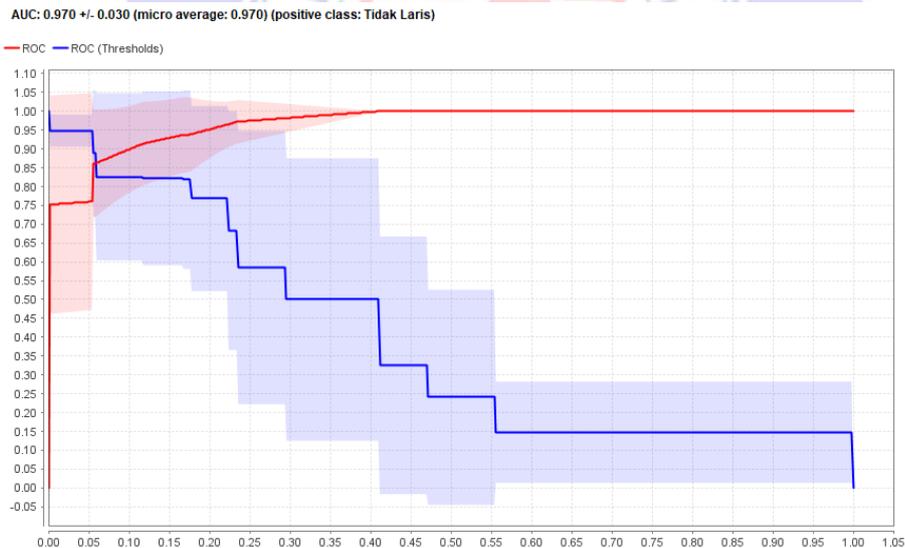
	true Laris	true Tidak Laris	class precision
pred. Laris	172	6	96.63%
pred. Tidak Laris	1	33	97.06%
class recall	99.42%	84.62%	

Sumber: Penelitian Mandiri (2024)

**Gambar III. 23.**  
Nilai *Recall* pada Algoritma C4.5

Pada **Gambar III.23** terlihat bahwa hasil nilai recall dari *Decision Tree* sebesar 85.17% dengan  $\pm 16.28\%$  dan *micro average* 84.62% dengan kasus kelas “Tidak Laris” jadi *positive*. Dari ketiga gambar diatas, dapat dilihat distribusi hasil prediksi dan data sebenarnya:

- 1) **True Laris** (data sebenarnya "Laris" yang diprediksi benar) sebanyak 172
- 2) **True Tidak Laris** (data sebenarnya "Tidak Laris" yang diprediksi benar) sebanyak 33
- 3) **False Laris** (data sebenarnya "Tidak Laris" yang diprediksi sebagai "Laris") sebanyak 6
- 4) **False Tidak Laris** (data sebenarnya "Laris" yang diprediksi sebagai "Tidak Laris") sebanyak 1



Sumber: Penelitian Mandiri (2024)

**Gambar III. 24.**  
Nilai *Area Under Curve* (AUC) pada Algoritma C4.5

Pada **Gambar III.24** terlihat bahwa hasil *Area Under Curve* dari *Decision Tree* sebesar 0.970 dengan  $\pm 0.030$ , termasuk kedalam kategori sangat baik.

## BAB IV

### KESIMPULAN

#### 4.1. Kesimpulan

Berdasarkan hasil penelitian, beberapa kesimpulan yang dapat ditarik adalah sebagai berikut:

1. Algoritma C4.5 efektif dalam menganalisis dan mengklasifikasikan data penjualan rokok, menghasilkan model yang mampu mengidentifikasi penjualan rokok terlaris dengan akurasi yang baik.
2. Penerapan metode ini di CV Jaya Berkah Mas memberikan wawasan yang bermanfaat bagi perusahaan dalam merencanakan strategi penjualan yang lebih efektif dan efisien. Penelitian ini juga menunjukkan pentingnya proses pembersihan data dalam memastikan kualitas analisis data mining.
3. Hasil penelitian ini juga didapatkan akurasi yang lebih tinggi dari penelitian sebelumnya, dengan menggunakan metode Algoritma C4.5 dan data sebanyak 212 *entry* yang menghasilkan tingkat Akurasi 96.73%, Presisi 98.00%, *Recall* 85.17%, *Area Under Curver* (AUC) 0.970 yang merupakan kategori sangat baik.
4. Dapat menghasilkan atribut baru sebagai output pada pohon keputusan seperti rasa, kemasan rokok, aroma dan hisapan rokok yang dihasilkan dari membandingkan penelitian terdahulu yang dilakukan oleh (Marliani et al., 2023) dan (Ridho & Rusda, 2024).
5. Dari hasil penelitian ini terdapat beberapa rekomendasi untuk penjualan rokok CV Jaya Berkah Mas kedepannya, yaitu:
  - a. Produksi rokok dengan rasa teh sebaiknya dikurangi
  - b. Produksi rokok dengan rasa manis, kemasan plastik, dan aroma lembut sebaiknya dikurangi

- c. Produksi rokok dengan rasa wangi cengkeh dan hisapan berat sebaiknya dikurangi.

## DAFTAR PUSTAKA

- Anggi Trifani, Agus Perdana Windarto, & Hendry Qurniawan. (2022). Penerapan Data Mining Klasifikasi C4.5 dalam Menentukan Tingkat Stres Mahasiswa Akhir. *Jurnal Riset Rumpun Ilmu Teknik*, 1(2), 92. <https://doi.org/10.55606/jurritek.v1i2.414>
- Arhami, M., & Nasir, M. (2020). *Data Mining.pdf* (I. R. Utami (ed.); 1st ed.). ANDI (IKAPI).
- Fauziningrum, E., & Suryaningsih, E. I. (2021). *Penerapan Data Mining Metode Decision Tree Untuk Mengukur Penguasaan Bahasa Inggris Maritim*. CV. PUSTAKA STIMAR AMNI SEMARANG.
- Leonardi, M., Emilda, R., Katrin, I., & Yulianto, A. (2021). Prediksi Penjualan Produk Rokok Pada PT. Indomarco Prismatama Menggunakan Algoritma C4.5. *Paradigma - Jurnal Komputer Dan Informatika*, 23(2), 182–190. <https://doi.org/10.31294/p.v23i2.11151>
- Marliani, T., Arini, D. P., & Santoso, I. (2023). Analisa Standar Produk Sabun Pembersih Baju. *Jurnal Komputer Dan Informatika*, 7(2), 115–123.
- Musa, D. M., Sakti, D., Shantiony, K. A., Zega, S. K. P., Hamzah, S., Zega, Y. J., & Lubis, B. O. (2024). Penerapan Data Mining Untuk Klasifikasi Data Penjualan Pakan Ternak Terlaris Dengan Algoritma C4.5. *Jurnal Teknologi Informatika Dan Komputer*, 10(1), 168–182. <https://doi.org/10.37012/jtik.v10i1.1985>
- Ovan, & Saputra, A. (2020). *Aplikasi Uji Validitas dan Reliabilitas Instrumen Penelitian Berbasis Web.pdf*. Yayasan Ahmar Cendekia Indonesia.
- Pratiwi, D. A., Awangga, R. M., & Setyawan, M. Y. H. (2020). Seleksi Calon Kelulusan Tepat Waktu Mahasiswa Teknik Informatika Menggunakan Metode Naive Bayes. In R. M. Awangga (Ed.), *Kreatif* (1st ed.). Kreatif Industri Nusantara.
- Qomariah, N., Pardede, A. J., Kamil, I., Asalia, N., & Nurbayani, S. (2024). Bagaimana Kualitas Produk, Harga Dan Promosi Meningkatkan Kepuasan Pelanggan Pada Industri Rokok Golden Taste 88. *BUDGETING: Journal of Business, Management and Accounting*, 5(2), 553. <https://doi.org/10.31539/budgeting.v5i2.7888>
- Ridho, S., & Rusda, D. (2024). Analisis Preferensi Konsumen dalam Memilih Produk Hortikultura Menggunakan Metode Algoritma C45 dan Naive Bayes. *Emitor: Jurnal Teknik Elektro*, 66–77. <https://doi.org/10.23917/emitor.v24i1.2401>
- Shaumi, A. R., Ali, M. F., & Simbolon, M. T. A. M. (2022). Penerapan Data Mining Menggunakan Metode Teknik Classification Untuk Melihat Penerapan Data Mining Menggunakan Metode Teknik Classification Untuk. *JUKI: Jurnal Komputer Dan Informatika*, 4(2), 174. [www.pajak.go.id](http://www.pajak.go.id)
- Sugara, B., Adidarma, D., & Budilaksono, S. (2019). Perbandingan Akurasi Algoritma C4.5 dan Naive Bayes untuk Deteksi Dini Gangguan Autisme pada Anak. *Jurnal IKRA-ITH Informatika*, 3(1), 122.
- Suriani, U. (2023). Penerapan Data Mining untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Decision Tree C4.5. *Journalcisa*, 3(2), 56. <https://doi.org/1051519/journalcisa.v4i2.393>
- Triawan, B., Lubis, I., & Kadim, Lina, Arliana, N. (2024). *Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper*. 3(November), 112–120.
- Tukino, T. (2019). Penerapan Algoritma C4.5 Untuk Memprediksi Keuntungan Pada PT SMOE Indonesia. *Jurnal Sistem Informasi Bisnis*, 9(1), 41.

<https://doi.org/10.21456/vol9iss1pp39-46>

Umradi, & Wijaya, H. (2020). *Analisis Data Kualitatif Teori Konsep dalam Penelitian Pendidikan.pdf*.

Wahyudi, M. D. (2023). Penerapan Data Mining Dengan Algoritma C4. 5 Dalam Prediksi Penjualan Buku. *Jurnal Teknorama (Informatika Dan ...*, 1(1), 1–6. <https://jurnal.stikomelrahma.ac.id/index.php/teknorama/article/view/1%0Ahttps://jurnal.stikomelrahma.ac.id/index.php/teknorama/article/download/1/1>

Zai, C. (2022). Implementasi Data Mining Sebagai Pengolahan Data. *Jurnal Teknoinfo*, 2(1), 6. <http://portaldata.org/index.php/portaldata/article/view/107>



# DAFTAR RIWAYAT HIDUP

## DAFTAR RIWAYAT HIDUP

### I. Biodata Mahasiswa

NIM : 19236055  
Nama Lengkap : Sebastianus Aditya Narendra  
Tempat/Tanggal Lahir : Tangerang Selatan, 18 Desember 2002  
Alamat Lengkap : Jl. Pondok Pakulonan Blok H6 No.3, RT.04/RW.04

### II. Pendidikan

1. SD Negeri Pakualam 01 Tangerang Selatan, Lulusan Tahun 2014
2. SMP Negeri 15 Tangerang Selatan, Lulusan Tahun 2017
3. SMK Negeri 6 Tangerang Selatan, Lulusan Tahun 2020

### III. Riwayat Pengalaman Berorganisasi/Pekerjaan

1. Praktik Kerja Lapangan di Rumah Belajar TOP SCORE
2. Menteri Informasi dan Komunikasi BEM UBSI BANYUMAS Periode 2022
3. Praketa Kopi Purwokerto



Jakarta, 04 Juli 2024

Sebastianus Aditya Narendra

# DAFTAR RIWAYAT HIDUP

## DAFTAR RIWAYAT HIDUP

### I. Biodata Mahasiswa

NIM : 19236049  
Nama : Pradita Aldi Setiawan  
Tempat/Tanggal Lahir : Purwokerto, 27 Januari 2002  
Alamat Lengkap : Jl. Kober Gg Mayang Rt 05/Rw 05,  
Kecamatan Purwokerto Barat 53132

### II. Pendidikan

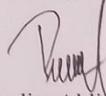
1. SD Negeri 2 Purwokerto Lor, Lulus Tahun 2014
2. SMP Negeri 3 Purwokerto, Lulus Tahun 2017
3. SMK Kesatrian Purwokerto, Lulus Tahun 2020

### III. Riwayat Pengalaman Berorganisasi/Pekerjaan

1. Palang Merah Remaja SMK
2. Himpunan Mahasiswa Sistem Informasi (HIMSI) UBSI



Jakarta, 04 Juli 2024

  
Pradita Aldi Setiawan

# LEMBAR KONSULTASI BIMBINGAN

	<b>LEMBAR KONSULTASI SKRIPSI</b>
	<b>UNIVERSITAS BINA SARANA INFORMATIKA</b>

NIM : 19236055  
Nama Lengkap : Sebastianus Aditya Narendra  
Dosen Pembimbing I : Syaiful Anwar, M.Kom  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing I
1	08 April 2024	Pengajuan Judul (online)	
2	22 April 2024	Pengajuan BAB I dan BAB II (online)	
3	04 Mei 2024	Revisi BAB I (online)	
4	18 Mei 2024	Konsultasi BAB III (online)	
5	31 Mei 2024	Pengajuan BAB III	
6	26 Juni 2024	Revisi Data dan BAB III	
7	27 Juni 2024	Penyesuaian Data dan BAB III	
8	03 Juli 2024	ACC Keseluruhan	

Catatan untuk Dosen Pembimbing.

Bimbingan Skripsi

Dimulai pada tanggal : 03 April 2024  
Diakhiri pada tanggal : 03 Juli 2024  
Jumlah pertemuan bimbingan : 8 (Delapan)

Disetujui oleh,  
Dosen Pembimbing I



(Syaiful Anwar, M.Kom)



**LEMBAR KONSULTASI SKRIPSI**  
**UNIVERSITAS BINA SARANA INFORMATIKA**

NIM : 19236055  
Nama Lengkap : Sebastianus Aditya Narendra  
Dosen Pembimbing I : Mohammad Haddiel Fuad, MTI  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing II
1	06 Mei 2024	Pengajuan dan ACC Judul (online)	<i>Fuad</i>
2	18 Mei 2024	Revisi BAB I dan BAB II (online)	<i>Fuad</i>
3	27 Juni 2024	Revisi Format Penulisan	<i>Fuad</i>
4	28 Juni 2024	Konsultasi BAB III (online)	<i>Fuad</i>
5	01 Juli 2024	Pengajuan BAB III	<i>Fuad</i>
6	02 Juli 2024	Revisi Data dan BAB III	<i>Fuad</i>
7	03 Juli 2024	Penyesuaian Data dan BAB III	<i>Fuad</i>
8	03 Juli 2024	ACC Keseluruhan	<i>Fuad</i>

Catatan untuk Dosen Pembimbing,

Bimbingan Skripsi

Dimulai pada tanggal : 06 Mei 2024

Diakhiri pada tanggal : 03 Juli 2024

Jumlah pertemuan bimbingan : 8 (Delapan)

Disetujui oleh,  
Dosen Pembimbing II

(Mohammad Haddiel Fuad, MTI)



**LEMBAR KONSULTASI SKRIPSI**  
**UNIVERSITAS BINA SARANA INFORMATIKA**

NIM : 19236049  
Nama Lengkap : Pradita Aldi Setiawan  
Dosen Pembimbing I : Syaiful Anwar, M.Kom  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing I
1	08 April 2024	Pengajuan Judul (online)	
2	22 April 2024	Pengajuan BAB I dan BAB II (online)	
3	04 Mei 2024	Revisi BAB I (online)	
4	18 Mei 2024	Konsultasi BAB III (online)	
5	31 Mei 2024	Pengajuan BAB III	
6	26 Juni 2024	Revisi Data dan BAB III	
7	27 Juni 2024	Penyesuaian Data dan BAB III	
8	03 Juli 2024	ACC Keseluruhan	

Catatan untuk Dosen Pembimbing.

Bimbingan Skripsi

Dimulai pada tanggal : 03 April 2024

Diakhiri pada tanggal : 03 Juli 2024

Jumlah pertemuan bimbingan : 8 (Delapan)

Disetujui oleh,  
Dosen Pembimbing I

(Syaiful Anwar, M.Kom)



## LEMBAR KONSULTASI SKRIPSI

UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19236049  
Nama Lengkap : Pradita Aldi Setiawan  
Dosen Pembimbing I : Mohammad Haddiel Fuad, MTI  
Judul Skripsi : Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada Cv Jaya Berkah Mas

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing II
1	06 Mei 2024	Pengajuan dan ACC Judul (online)	<i>[Signature]</i>
2	18 Mei 2024	Revisi BAB I dan BAB II (online)	<i>[Signature]</i>
3	27 Juni 2024	Revisi Format Penulisan	<i>[Signature]</i>
4	28 Juni 2024	Konsultasi BAB III (online)	<i>[Signature]</i>
5	01 Juli 2024	Pengajuan BAB III	<i>[Signature]</i>
6	02 Juli 2024	Revisi Data dan BAB III	<i>[Signature]</i>
7	03 Juli 2024	Penyesuaian Data dan BAB III	<i>[Signature]</i>
8	03 Juli 2024	ACC Keseluruhan	<i>[Signature]</i>

Catatan untuk Dosen Pembimbing.

Bimbingan Skripsi

Dimulai pada tanggal : 06 Mei 2024

Diakhiri pada tanggal : 03 Juli 2024

Jumlah pertemuan bimbingan : 8 (Delapan)

Disetujui oleh,  
Dosen Pembimbing II

(Mohammad Haddiel Fuad, MTI)

# SURAT KETERANGAN RISET



**CV JAYA BERKAH MAS**  
**Jl. Mustari No. 17 Dukuh Bongok RT 03 RW 09**  
**Penaruban - Purbalingga**  
**Telp. (081329900200)**

## SURAT KETERANGAN

Yang bertanda tangan di bawah ini :

Nama : Adi Wardoyo, S. Kom  
Jabatan : Kepala Gudang CV Jaya Berkah Mas

Dengan ini menerangkan bahwa, yang tersebut di bawah ini :

Nama : Sebastianus Aditya Narendra  
NIM : 19236055  
Alamat : Jl. Raya PUK, Tumiayang, Kebasen, Rt 02/Rw 02

Adalah Benar telah melakukan riset pada CV Jaya Berkah Mas sejak 10 April 2024 sampai dengan 10 Juni 2024, dan yang bersangkutan telah melakukan tugasnya dengan baik dan penuh tanggung jawab.

Demikian surat keterangan ini dibuat dengan benar, untuk dapat dipergunakan sebagaimana mestinya.

Purwokerto, 21 Juni 2024

  
 **CV JAYA BERKAH MAS**

Adi Wardoyo, S. Kom

Kepala Gudang CV Jaya Berkah Mas



**CV JAYA BERKAH MAS**  
**Jl. Mustari No. 17 Dukuh Bongok RT 03 RW 09**  
**Penaruban - Purbalingga**  
**Telp. (081329900200)**

---

**SURAT KETERANGAN**

Yang bertanda tangan di bawah ini :

Nama : Adi Wardoyo, S. Kom  
Jabatan : Kepala Gudang CV Jaya Berkah Mas

Dengan ini menerangkan bahwa, yang tersebut di bawah ini :

Nama : Pradita Aldi Setiawan  
NIM : 19236049  
Alamat : Jl. Kober Gg Mayang RT 05 / RW 05 Kec. Purwokerto Barat

Adalah Benar telah melakukan riset pada CV Jaya Berkah Mas sejak 10 April 2024 sampai dengan 10 Juni 2024, dan yang bersangkutan telah melakukan tugasnya dengan baik dan penuh tanggung jawab.

Demikian surat keterangan ini dibuat dengan benar, untuk dapat dipergunakan sebagaimana mestinya.

Purwokerto, 21 Juni 2024

  
**CV. JAYA BERKAH MAS**

Adi Wardoyo, S. Kom  
Kepala Gudang CV Jaya Berkah Mas

## LAMPIRAN – LAMPIRAN

### Lampiran 1. Produk Rokok





**Lampiran 3. Personil dan Jadwal Tim Jaya Berkah Purbalingga**

**PERSONIL TIM PURBALINGGA**

No.	NAMA SALES	AREA KECAMATAN	KETERANGAN
1	KHOLIL	BOBOTSARI DAN KERTANEGARA	
2	DIDIN	SUMBANG, KEMBARAN DAN BATURADEN	
3	RAFI	BELIK, KARANGREJA, KARANGJAMBU DAN MREBET	
4	BADRUN	KUTASARI PADAMARA DAN PBG KOTA	
5	BASUKI	KALIMANAH DAN KEMANGKON	JOINT TGL 16 APRIL 2024
6	FAJAR	KARANGMONCOL DAN REMBANG	JOINT TGL 17 APRIL 2024
7	YUDI	BUKATEJA DAN KEJOBONG	JOINT TGL 19 APRIL 2024
8	REDIAN	BOJONGSARI KALIGONDANG DAN PANGADEGAN	JOINT TGL 22 APRIL 2024

**JADWAL TIM PROMOTION**

SELASA 16	RABU 17	KAMIS 18	JUMAT 19	SABTU 20
PS KALIGONDANG	PS KEBUTUH	PS BANDINGAN	OFF	PS SOKAWERA
PS TAMBARAN	PS HWN PURWONEGORO	PS SEGAMAS		PS SOKARAJA
PS SINDURAJA	PS PUCANG	PS HWN PURBALINGGA		PABRIK KAYU KALIBAGOR
PS KEJOBONG				PS BANYUMAS
MINGGU 21	SENIN 22	SELASA 23	RABU 24	KAMIS 25
PS WAGE PHT	PS SEGAMAS	PS TAMBAKSOGRA	PS BOBOTSARI	PS BANDINGAN
TERMINAL 1 DAN 2 BATURADEN	PS BANDINGAN	PS LARANGAN	PS BELIK	PS SEGAMAS
KOMPLEK KAMPUS DAN PERTOKAN UNSOED	PS HEWAN PURBALINGGA	PS CERME	PS SAYUR PRATIN	PS HWN PURBALINGGA
JUMAT 26	SABTU 27	MINGGU 28	SENIN 29	SELASA 30
	PS KARANGREJA			



## BUKTI HASIL PENGECEKAN PLAGIARISME

Analisa Penjualan Rokok Dengan Metode Klasifikasi Menggunakan Formula Algoritma C4.5 Pada CV Jaya Berkah Mas

### ORIGINALITY REPORT

<b>23%</b> SIMILARITY INDEX	<b>21%</b> INTERNET SOURCES	<b>15%</b> PUBLICATIONS	<b>9%</b> STUDENT PAPERS
--------------------------------	--------------------------------	----------------------------	-----------------------------

### PRIMARY SOURCES

<b>1</b>	<b>docplayer.info</b> Internet Source	<b>3%</b>
<b>2</b>	<b>repository.nusamandiri.ac.id</b> Internet Source	<b>1%</b>
<b>3</b>	<b>journal.thamrin.ac.id</b> Internet Source	<b>1%</b>
<b>4</b>	<b>ejournal.bsi.ac.id</b> Internet Source	<b>1%</b>
<b>5</b>	<b>ejournal.upbatam.ac.id</b> Internet Source	<b>1%</b>
<b>6</b>	<b>eiurnal.stie-trianandra.ac.id</b>	<b>1</b>

UNIVERSITAS