

PENERAPAN DATA MINING UNTUK PREDIKSI PENJUALAN PRODUK ELEKTRONIK TERLARIS MENGGUNAKAN METODE DECISION TREE

Aida Roswita Suta¹, Yoseph Pius Kurniawan Kelen², Krisantus J.T. Seran³
^{1,2,3}Universitas Timor

^{1,2,3}Program Studi Teknologi Informasi Fakultas Pertanian Sains dan Kesehatan Universitas Timor
KM. 9 Kelurahan Sasi Kefamenanu

e-mail: ¹aidasuta512@gmail.com, ²yosepkelen@unimor.ac.id, ³krisantusteyseran@unimor.ac.id

Abstrak

Dalam dunia bisnis, khususnya dalam industri elektronik, prediksi penjualan menjadi faktor penting dalam menentukan strategi pemasaran dan manajemen stok. Data mining merupakan teknik yang dapat digunakan untuk menggali pola dan tren dari data historis penjualan guna menghasilkan prediksi yang lebih akurat. Penelitian ini bertujuan untuk menerapkan metode Decision Tree dalam memprediksi produk elektronik terlaris berdasarkan data transaksi penjualan. Metode Decision Tree dipilih karena kemampuannya dalam mengelompokkan data dan menghasilkan model yang mudah dipahami. Data yang digunakan dalam penelitian ini mencakup variabel seperti kategori produk, harga, jumlah penjualan, musim pembelian, dan faktor lainnya yang berkontribusi terhadap keputusan pembelian pelanggan. Proses analisis meliputi pengumpulan data, pembersihan data, eksplorasi pola, serta pembuatan dan evaluasi model Decision Tree. Hasil penelitian menunjukkan bahwa metode Decision Tree mampu mengidentifikasi faktor-faktor utama yang mempengaruhi penjualan produk elektronik dan memberikan prediksi yang cukup akurat. Model yang dihasilkan dapat membantu perusahaan dalam mengoptimalkan strategi pemasaran dan manajemen stok guna meningkatkan profitabilitas. Dengan penerapan data mining, khususnya metode Decision Tree, perusahaan dapat mengambil keputusan yang lebih berbasis data dan efektif dalam menghadapi persaingan pasar.

Kata Kunci: Data Mining, Prediksi, Produk Elektronik, Decision Tree

Abstract

In the business world, especially in the electronics industry, sales predictions are an important factor in determining marketing strategies and stock management. Data mining is a technique that can be used to explore patterns and trends from historical sales data to produce more accurate predictions. This research aims to apply the Decision Tree method in predicting best-selling electronic products based on sales transaction data. The Decision Tree method was chosen because of its ability to group data and produce a model that is easy to understand. The data used in this research includes variables such as product category, price, number of sales, purchasing season, and other factors that contribute to customer purchasing decisions. The analysis process includes data collection, data cleaning, pattern exploration, and decision tree model creation and evaluation. The research results show that the Decision Tree method is able to identify the main factors that influence sales of electronic products and provide fairly accurate predictions. The resulting model can help companies optimize marketing and stock management strategies to increase profitability. By applying data mining, especially the Decision Tree method, companies can make more data-based and effective decisions in facing market competition.

Keywords: Data Mining, Prediction, Electronic Products, Decision Tree

1. PENDAHULUAN

Prediksi atau peramalan penjualan (*forecasting*) merupakan suatu perhitungan untuk meramalkan keadaan dimasa mendatang dengan melakukan pengujian terhadap keadaan di masa lalu. Salahsatu dari fungsi prediksi adalah untuk membantu pemilik perusahaan / toko dalam pengambilan keputusan dan menentukan jumlah barang yang harus disediakan oleh perusahaan / toko tersebut. Selain itu, prediksi dapat membantu pihak perusahaan dalam penyediaan stok barang, karena prediksi ini dapat memberikan output terbaik sehingga diharapkan resiko kesalahan yang disebabkan oleh kesalahan perencanaan dapat ditekan seminimal mungkin. Prediksi biasanya digunakan untuk menemukan informasi dari data yang besar sehingga diperlukan data mining[1].

Data mining merupakan teknik pembelajaran mesin, pengenalan pola, database, statistic dan visualisasi yang digunakan untuk penanganan masalah pengambilan informasi dari penyimpanan database yang besar. Data mining dapat digunakan untuk menggali informasi dari data yang besar sehingga didapatkan informasi dalam memprediksi penjualan barang elektronik. Dalam data mining terdapat banyak teknik dalam pengerjaannya untuk menemukan pola atau informasi yang tersembunyi diantaranya adalah klasterisasi (*clustering*), regresi (*regression*), asosiasi (*association*), dan klasifikasi (*classification*). Metode- metode yang terdapat dalam data mining untuk prediksi diantaranya yaitu metode *Decision Tree*[2]. Metode ini digunakan untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Penggunaan metode *Decision Tree* sebagai prediksi data sudah banyak digunakan, diantaranya algoritma *Decision Tree* dalam aktivitas manusia dalam pembelajaran online. *Decision Tree* sebagai metode untuk klasifikasi terhadap data untuk memprediksi barang-barang elektronik terlaris per-tahunnya [3].

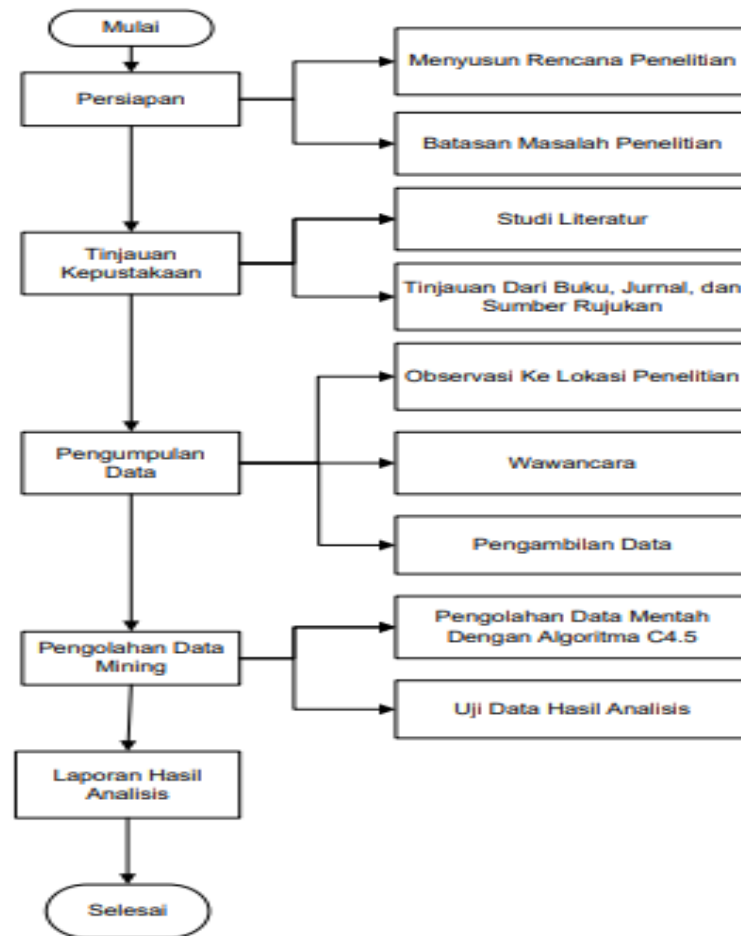
Penelitian ini menggunakan dukungan data disebuah toko elektronik yang berada di kefamenanu ("Aries Elektronik"). Aries Elektronik merupakan salah satu toko elektronik yang bergerak dibidang perdagangan yang menyediakan berbagai macam produk elektronik dan furniture. Aries Elektronik saat ini telah menggunakan sistem informasi penjualan untuk mendukung kinerja pihak toko maupun layanan informasi penjualan. Toko ini juga telah menerapkan sistem penjualan secara online dalam mempromosikan penjualannya. Produk yang dijual diantaranya : Televisi (TV), Air Condisioner (AC), Mesin Cuci dan produk elektronik lainnya. Dilihat dari banyaknya permintaan konsumen akan produk elektronik, maka dibutuhkan prediksi untuk penjualan elektronik yang paling banyak diminta oleh konsumen atau pelanggan. Data tersebut kemudian diolah menggunakan metode *Decision Tree*[4]. Data hasil olahan ini menggunakan metode *Decision Tree* yang akan disajikan kepada user[5]. Data ini diharapkan dapat membantu user dalam memprediksi penjualan barang-barang elektronik terlaris dan prediksi ini bertujuan untuk mempermudah bagian penyedia stok barang pada toko Aries Elektronik dalam melakukan perencanaan penyediaan stok barang serta memberitahu pihak toko tentang produk-produk paling banyak dibeli oleh konsumen[6].

Dalam penelitian ini algoritma yang dipakai dalam metode *Decission Tree* adalah C4.5 yang merupakan sebuah algoritma pembelajaran mesin yang digunakan untuk membangun model prediksi berdasarkan data yang tersedia. Algoritma ini termasuk dalam kelas algoritma *Decission Tree* yang berarti bahwa ia membuat keputusan berdasarkan serangkaian pertanyaan yang terstruktur dalam bentuk pohon keputusan. C4.5 diperkenalkan oleh J.Ross Quinian pada tahun 1993 dan algoritma ini memiliki beberapa keunggulan dibandingkan algoritma *Decission Tree* lainnya, seperti kemampuan untuk menangani data numerik dan data kategorik serta kemampuan untuk mempertahankan akurasi model meskipun terjadi perubahan pada data. Algoritma ini sering digunakan dalam aplikasi machine learning untuk memprediksi keputusan atau klasifikasi suatu data berdasarkan fitur-fitur yang ada. Didalam algoritma C4.5 terdapat 2 rumus yaitu mencari nilai entropy dan untuk mencari nilai gain[8]. Data olahan yang di ambil mencapai 500 data mulai dari Januari 2020 sampai Desember 2023.

2. METODE PENELITIAN

2.1. Metodologi Penelitian

Gambar berikut merupakan ilustrasi yang memperlihatkan tahapan penelitian yang dilakukan oleh peneliti adalah sebagai berikut[9] :



Gambar 1. Tahapan Penelitian

Tahapan penelitian yang dituangkan dalam diagram alir ini menggambarkan penelitian secara keseluruhan. Tahapan yang ditempuh yaitu :

1. Persiapan. Tahap ini merupakan tahapan subjek (populasi) toko Aries Elektronik Kefamenanu. Objek yang diambil yaitu produk elektronik, batasan dan menyusun rencana penelitian.
2. Tinjauan Kepustakaan. Dalam tinjauan kepustakaan dilakukan telaah dan studi literatur mengenai prediksi penjualan dan yang berhubungan.
3. Pengumpulan Data. Pengumpulan data dilakukan dengan cara wawancara kepada bagian *accounting*, observasi dan dokumen.
4. Pengolahan Data Mining. Data yang dikumpulkan diolah sesuai dengan tahapan *knowledge discovery in database (KDD)*.

Hasil. Pada tahapan ini menjelaskan hasil dari proses dari data mining yang dilakukan dengan menggunakan metode *Decision Tree*.

2.2. Data Mining

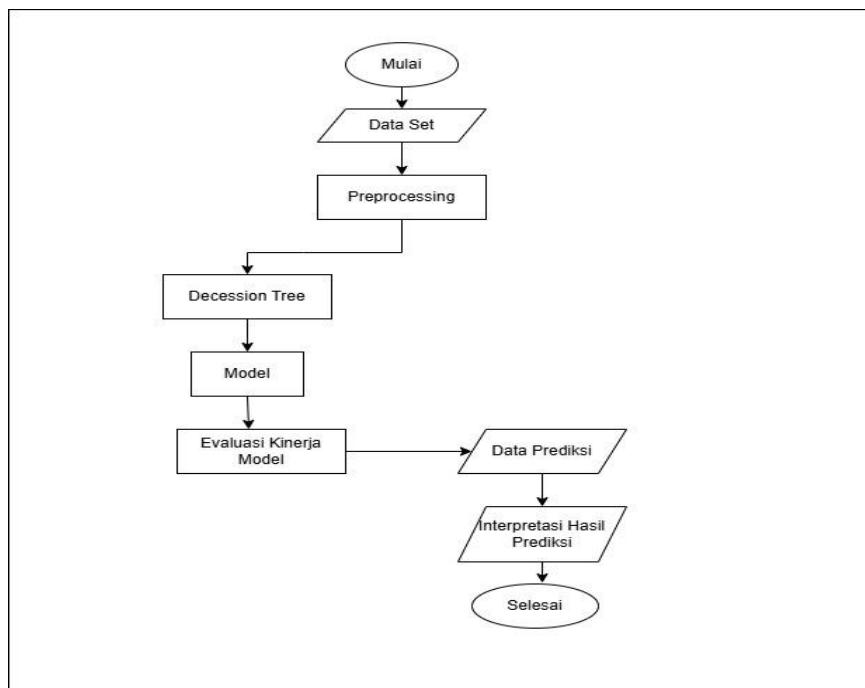
Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstraksi pengetahuan (*knowlndge*) secara otomatis. Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.

Data mining juga merupakan suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning*, untuk memngekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait.

Berdasarkan definisi-definisi di atas tentang data mining dapat disimpulkan bahwa data mining adalah sebuah proses pencarian secara otomatis untuk menemukan pola atau model dari suatu database yang besar[10].

2.3. Prediksi Penjualan

Prediksi/*forecasting* adalah menentukan jumlah kebutuhan bulan mendatang terkait dengan dukungan data historis (*historical data*) atau serangkaian waktu/periode yang dianalisis sehingga dapat diperhitungkan untuk memprediksi jumlah kebutuhan pada bulan mendatang. Prediksi juga dapat digunakan dalam pengklasifikasian, tidak hanya untuk memprediksi *time series* karena sifatnya yang bisa menghasilkan *class* berdasarkan atribut yang ada[11]. Berikut adalah *flowchart* model dalam prediksi data :



Gambar 2. Flowchart Dalam Prediksi Data

2.4. Produk Elektronik

Aries Elektronik merupakan salah satu toko elektronik yang berada di kota kefamenanu, dimana toko tersebut banyak diminati oleh para konsumen karena tersedia berbagai macam produk elektronik seperti : Air Condisioner (AC), Kulkas, Televisi (TV), Kipas Angin, Laptop, Parabola, dan masih banyak lagi barang-barang elektronik lainnya. Sehingga perlu adanya prediksi untuk membantu pemilik toko dalam memperkirakan keluar masuknya stok barang yang banyak diminta oleh para konsumen/pembeli.

2.5. Decision Tree

Algoritma *Decision Tree* dipahami sebagai teknik model prediksi yang dapat digunakan untuk klasifikasi dan prediksi tugas. *Decision Tree* menggunakan teknik “membagi dan menaklukkan” untuk membagi ruang pencarian masalah menjadi himpunan masalah. Proses pada *Decision Tree* adalah mengubah bentuk data tabel menjadi sebuah model *Tree*. Model *Tree* akan menghasilkan rule dan disederhanakan. Hal yang harus dilakukan dalam metode decision tree adalah menghitung *entropy* dan *information gain*.

Berikut persamaan (1) dari rumus entropy yaitu :

$$\text{Entropy}(s) = \sum_{j=1}^k -p_j \log_2 p_j$$

Dimana:

- S adalah himpunan (data set) kasus
- k adalah banyaknya partisi S
- p_j adalah probabilitas yang didapatkan dari $\text{sum}(Y_a)$ dibagi total kasus

Sedangkan untuk persamaan (2) rumus information gain yaitu:

$$\text{Gain}(A) = \text{Entropi}(s) - \sum_{i=1}^k \frac{|s_i|}{|s|} \times \text{Entropi}(s_i)$$

Dimana :

- S = ruang (data) sampel yang digunakan untuk training
- A = atribut
- $|s_i|$ = jumlah sampel untuk nilai V
- $|s|$ = jumlah seluruh sampel data
- Entropi (s_i) = entropi untuk sampel-sampel yang memiliki nilai i

Metode *Decision Tree* adalah salah satu metode yang digunakan untuk keperluan klasifikasi pada teknik data mining. Metode ini berfungsi untuk menerjemahkan fakta dalam lingkup sangatt besar menjadi pohon keputusan yang menyatakan dan menampilkan suatu aturan (*rule*). Pohon keputusan dapat juga digunakan untuk melakukan eksplorasi data, menentukan korelasi tersembunyi antara sekelompok calon yang akan menjadi variabel input dengan suatu variabel yang menjadi target. Data yang muncul dalam pohon keputusan umumnya dipresentasikan dalam suatu tabel yng didalamnya berisi atribut dan *record*. Konsep dari pohon keputusan adalah menerjemahkan suatu data menjadi pohon keputusan.



Gambar 3. Konsep *Decision Tree*

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan Data

Setelah melakukan wawancara dan observasi, Data sampel yang digunakan adalah data penjualan produk elektronik terlaris tahun 2020/2023 sebanyak 500 data. Adapun rincian data *training* dan data *testing* yang digunakan dalam penelitian ini yang dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1. Data Training

No	Merek	Jenis Barang	Tahun	Harga	Kualitas	Laku
1	Lenovo	Leptop	2020	Mahal	Sedang	Ya
2	Samsung	Leptop	2020	Murah	Rendah	Ya
3	Hp	Leptop	2020	Mahal	Sedang	Tidak
4	Dell	Leptop	2020	Murah	Sedang	Ya
5	Acer	Leptop	2020	Mahal	Tinggi	Ya
6	Toshiba	Leptop	2020	Murah	Rendah	Ya
7	Asus	Leptop	2020	Mahal	Tinggi	Tidak
8	Lg	Pc	2020	Murah	Rendah	Ya

9	Asus	Pc	2020	Mahal	Tinggi	Tidak
10	Dell	Pc	2020	Mahal	Sedang	Ya
...	...	...	...	...	...	...
500	Philip	Dvd Player	2023	Murah	Sedang	Tidak

Jadi, jumlah data *training* yang digunakan sebanyak 500 data barang elektronik dan rincian data *training* selengkapnya dapat dilihat pada Lampiran 4. Variabel atau atribut yang digunakan dalam penelitian ini yaitu data barang elektronik (merek, jenis barang, tahun, harga kualitas, dan laku). Adapun rincian data *testing* yang digunakan pada penelitian ini dapat dilihat pada Tabel 2.

Tabel 2. Data Testing

Atribut	Keterangan	Jumlah Data	Ya	Tidak	Entropy	Gain
Total		500	167	333	0.918961059	
Kualitas		500	167	333		0.084348324
	Rendah	192	99	93	0.999295444	
	Sedang	132	42	90	0.902393283	
	Tinggi	176	26	150	0.604123458	
Harga		500	167	333		0.088015286
	Murah	208	110	98	0.997597732	
	Mahal	292	57	235	0.712234787	
Tahun		500	167	333		0.001718863
	2020	125	44	81	0.935841553	
	2021	125	42	83	0.920939448	
	2022	125	44	81	0.935841553	
	2023	125	37	88	0.876346229	
Merek		500	167	333		0.054065907
	Lenovo	72	44	28	0.964078765	
	Samsung	72	14	58	0.710676854	
	Hp	68	26	42	0.959686894	
	Dell	72	17	55	0.788513161	
	Acer	72	27	45	0.954434003	
	Toshiba	72	18	54	0.811278124	
	Asus	72	21	51	0.870864469	

Berdasarkan Tabel di atas variabel atau atribut yang digunakan dalam penelitian ini yaitu data barang elektronik (merek, jenis barang, tahun, arga kualitas, dan laku). Dalam penelitian ini hanya menggunakan beberapa atribut karena sesuai dengan kriteria yang di gunakan oleh pihak toko dalam memprediksi produk elektronik terlaris.

3.2 Hasil Pengolahan Data

3.2.1 Import Data Microsoft Excel

Tahapan pertama yang dilakukan yaitu mengimport data dengan melihat nilai *gain* terbesar yang ada pada harga, kualitas, merek, dan tahun. Nilai terbesar disini adalah harga. Berikut merupakan hasil data yang digunakan untuk *Rapidminer*.

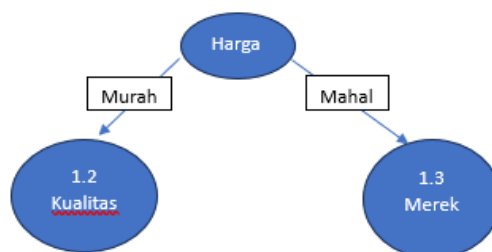
Tabel 3. Hasil Import Data

Atribut	Keterangan	Jumlah Data	Ya	Tidak	Entropy	Gain
Total		500	167	333	0.918961059	
Kualitas		500	167	333		0.084348324
	Rendah	192	99	93	0.999295444	
	Sedang	132	42	90	0.902393283	
	Tinggi	176	26	150	0.604123458	
Harga		500	167	333		0.088015286
	Murah	208	110	98	0.997597732	
	Mahal	292	57	235	0.712234787	
Tahun		500	167	333		0.001718863
	2020	125	44	81	0.935841553	
	2021	125	42	83	0.920939448	
	2022	125	44	81	0.935841553	
	2023	125	37	88	0.876346229	
Merek		500	167	333		0.054065907
	Lenovo	72	44	28	0.964078765	
	Samsung	72	14	58	0.710676854	
	Hp	68	26	42	0.959686894	
	Dell	72	17	55	0.788513161	
	Acer	72	27	45	0.954434003	
	Toshiba	72	18	54	0.811278124	
	Asus	72	21	51	0.870864469	

Berdasarkan Tabel di atas merupakan data yang di import sebanyak 500 data yang terdiri dari 4 atribut yaitu kuliatas, harga, tahun dan merek dengan keterangan (rendah, sedang dan tinggi) dan barang yang laku sebanyak 167 data dan tidak laku sebanyak 333 data dengan nilai *gain* terbesar yang ada pada atribut harga, kulitas, merek dan tahun.

3.2.2 Menghitung Nilai Entropy

Selanjutnya adalah melihat nilai entropy harga mahal dan murah. nilai entropy yang nilainya besar adalah nilai entropy murah dan akan menuju ke kualitas karena kualitas memiliki nilai terbesar ke 2. Sedangkan untuk nilai entropy mahal akan menuju ke merek karena merek merupakan nilai terbesar ke 3. Hasilnya akan seperti gambar di bawah ini :

**Gambar 4.** Tampilan Upload Harga

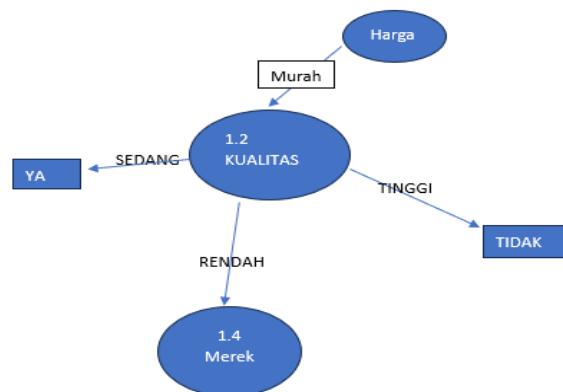
3.2.3 Algoritma C4.5 Berdasarkann Harga Murah Terhadap Kualitas

Tahap selanjutnya adalah membuat algoritma C4.5 berdasarkan harga murah terhadap kualitas. Berikut ini tampilan *upload* file data yang dapat dilihat pada gambar 4.5

Tabel 4. Tampilan *upload* data pada *microsoft Exel*

Atribut	Keterangan	Jumlah Data	Ya	Tidak	Entropy	Gain
Kualitas (Murah)		208	110	98	0.997597732	0.072439169
	Rendah	160	94	66	0.97779457	
	Sedang	20	12	8	0.970950594	
	Tinggi	28	4	24	0.591672779	

Dari tabel diatas nilai *entropy* terbesar yaitu nilai *entropy* rendah dan akan terhubung Kembali dengan merek. dapat di ambil gambar sebagai berikut :



Gambar 5. Tampilan *Upload* Kualitas

Nilai *entropy* tertinggi akan terhubung ke berikutnya. Kualitas tinggi panahnya menuju ke (tidak) karena jumlah Tidak lebih banyak dari jumlah (ya). Kualitas sedang (ya) karena jumlah (ya) lebih besar dari jumlah (tidak).

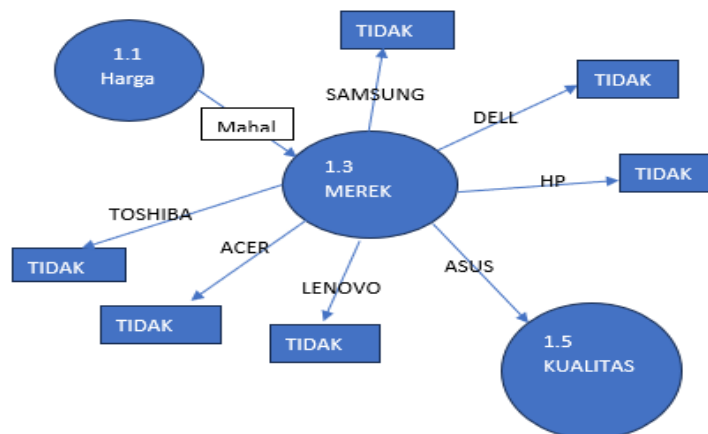
3.2.4 Algoritma C4.5 Berdasarkann Harga Mahal Terhadap Merek

Langkah selanjutnya adalah membuat algoritma C4.5 berdasarkan harga mahal terhadap merek yang akan ditampilkan pada tabel di bawah ini :

Tabel 5. Tampilan harga mahal terhadap merek

Atribut	Keterangan	Jumlah Data	Ya	Tidak	Entropy	Gain
Merek (Mahal)		292	57	235	0.712234787	0.044244247
	Lenovo	20	4	16	0.721928095	
	Samsung	72	14	58	0.710676854	
	Hp	28	2	26	0.371232327	
	Dell	32	8	24	0.811278124	
	Acer	48	11	37	0.776555785	
	Toshiba	44	2	42	0.266764988	
	Asus	48	16	32	0.918295834	

Nilai *entropy* tertinggi akan terhubung ke berikutnya. Dapat dilihat pada gambar 6.



Gambar 6. Tampilan Upload Merek

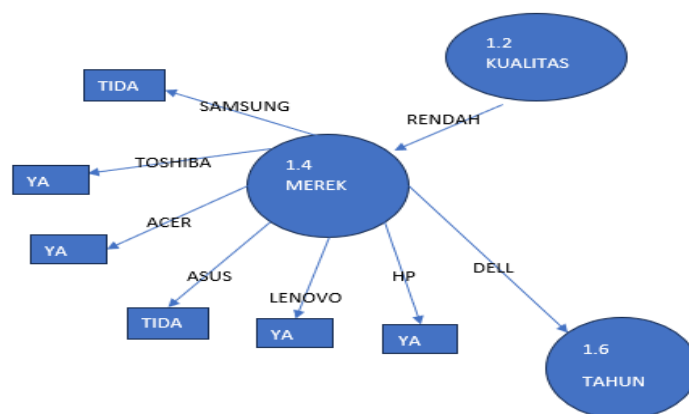
3.2.5 Algoritma C4.5 Berdasarkann Kualitas Rendah Terhadap Merek

Setelah mengupload dan menampilkan nilai *entropy* tertinggi dari harga mahal terhadap merek, selanjutnya buat algoritma c4.5 berdasarkan kualitas rendah terhadap merek. Berikut yang dapat dilihat pada tabel 6.

Tabel 6. Tampilan Kualitas Rendah Terhadap Merek

Atribut	Keterangan	Jumlah Data	Ya	Tidak	Entropy	Gain
Merek (Rendah)		192	99	93	0.999295444	0.171243162
	Lenovo	52	40	12	0.779349837	
	Samsung	32	5	27	0.625262405	
	Hp	20	12	8	0.970950594	
	Dell	20	9	11	0.992774454	
	Acer	24	16	8	0.918295834	
	Toshiba	20	12	8	0.970950594	
	Asus	24	5	19	0.738284866	

Nilai *entropy* tertinggi akan terhubung ke berikutnya. Dapat dilihat pada gambar 7.



Gambar 7. Tampilan Upload Merek

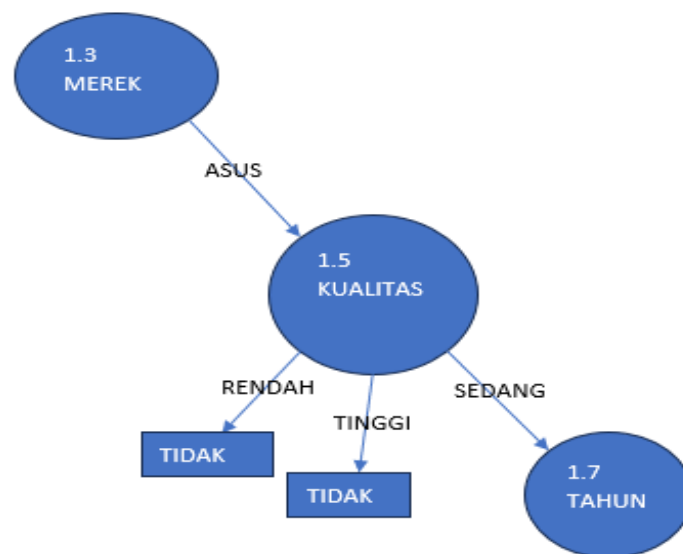
3.2.6 Algoritma C4.5 Berdasarkan Merek Asus Terhadap Kualitas

Langkah berikut adalah membuat algoritma c4.5 berdasarkan merek asus terhadap Kualitas. Dapat dilihat pada tabel di bawah ini.

Tabel 7. Tampilan Upload Merek Asus Terhadap Kualitas

Atribut	Keterangan	Jumlah Data	Ya	Tidak	Entropy	Gain
Kualitas (Asus)		72	21	51	0.870864469	0.041089657
	Rendah	24	5	19	0.738284866	
	Sedang	28	12	16	0.985228136	
	Tinggi	20	4	16	0.721928095	

Nilai *entropy* tertinggi akan terhubung ke berikutnya. Dapat dilihat pada gambar 8. Berikut ini :



Gambar 8. Nilai Entropy Kualitas

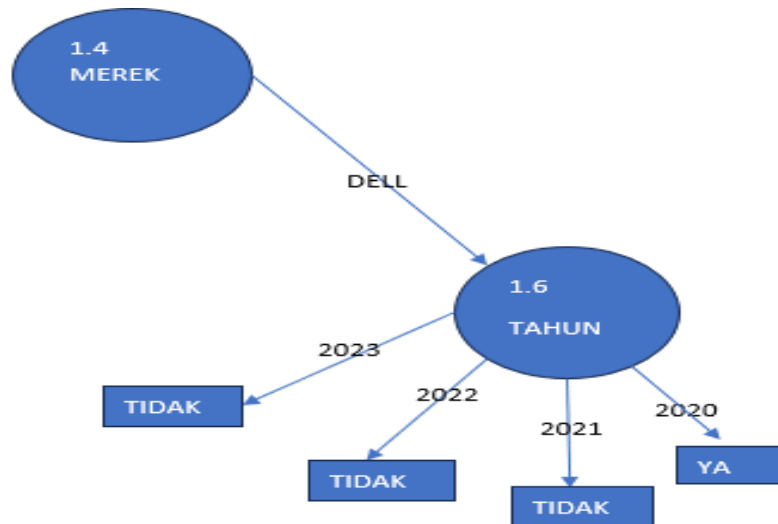
3.2.7 Algoritma C4.5 Berdasarkan Merek Dell Terhadap Tahun

Tahap selanjutnya adalah membuat algoritma C4.5 berdasarkan merek dell terhadap tahun. Berikut tampilannya dapat dilihat pada tabel 8.

Tabel 8. Tampilan Upload Merek Dell Terhadap Tahun

Atribut	Keterangan	Jumlah Data	Ya	Tidak	Entropy	Gain
Tahun (Dell)		72	17	55	0.788513161	0.193309613
	2020	18	5	13	0.852405179	
	2021	18	4	14	0.764204507	
	2022	18	4	14	0.764204507	
	2022	18	4	14	0.764204507	

Dari harga murah ke kualitas rendah kemudian lanjut ke merek, dari merek dell akan lanjut ke tahun. Tampilannya dapat dilihat pada gambar 9.



Gambar 9. Tampilan Upload Tahun

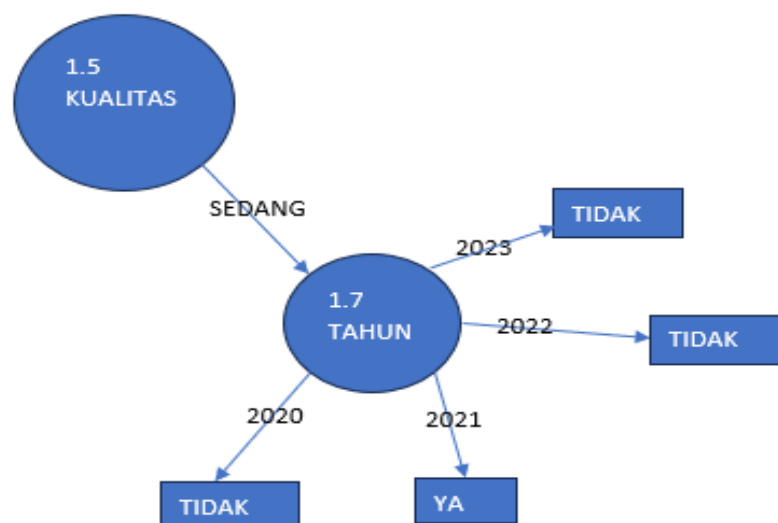
3.2.8 Algoritma C4.5 Berdasarkan Kualitas Sedang Terhadap Tahun

Langkah berikuta dalah membuat algoritma c4.5 berdasarkan kualitas sedang terhadap tahun. Dapat dilihat pada tabel di bawah ini.

Tabel 9. Tampilan Upload Kualitas Sedang Terhadap Tahun

Atribut	Keterangan	Jumlah Data	Ya	Tidak	Entropy	Gain
Tahun (Sedang)		132	42	90	0.902393283	0.213671407
	2020	33	11	22	0.918295834	
	2021	33	11	22	0.918295834	
	2022	33	11	22	0.918295834	
	2022	33	11	22	0.918295834	

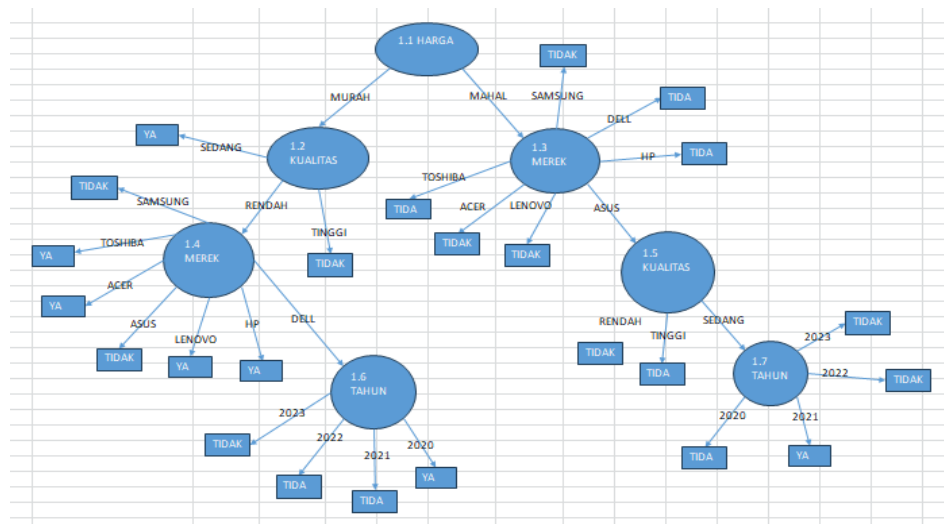
Setelah semua terhubung maka prosesnya akan selesai. Tampilannya dapat dilihat pada gambar 10.



Gambar 10. Tampilan Upload Tahun

3.2.9 Tampilan Pohon Keputusan Menggunakan Microsoft Excel

Setelah melakukan pengujian tahap demi tahap, maka akan muncul hasil dari algoritma C4.5 untuk membentuk pohon keputusan. Jadi sudah ditemukan ada beberapa pohon yang menjadi outputnya. Berikut tampilan poho keputusan menggunakan *microsoft excel* dapat dilihat pada gambar 11.



Gambar 11. Tampilan Pohon Keputusan

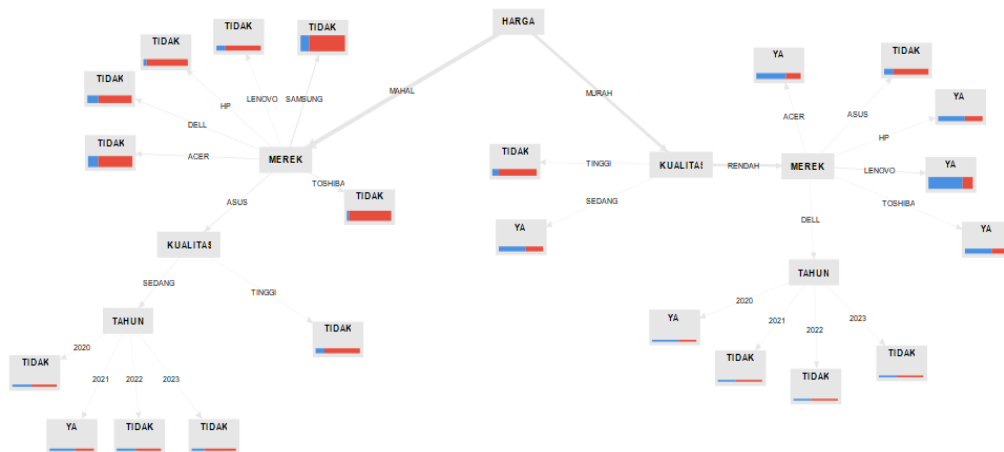
3.2.10 Perhitungan Menggunakan Rapidminer

Rapidminer merupakan salah satu *tools* yang dipakai dalam data mining. *Rapidminer* memudahkan penggunaanya dalam melakukan perhitungan data yang sangat banyak dengan menggunakan operator-operator. Langkah awal dalam membuat perhitungan menggunakan *rapidminer* adalah mengimport data. Disini data yang digunakan adalah sebanyak 500 data. Dapat dilihat pada tabel 10.

Tabel 10. Hasil Prediksi Menggunakan *Rapidminer*

Row No.	MERKEK	TAHUN	HARGA	KUALITAS	LAKU
1	LENOVO	2020	MURAH	RENDAH	YA
2	LENOVO	2020	MURAH	RENDAH	YA
3	LENOVO	2020	MAHAL	TINGGI	TIDAK
4	SAMSUNG	2020	MAHAL	RENDAH	YA
5	SAMSUNG	2020	MAHAL	SEDANG	TIDAK
6	SAMSUNG	2020	MAHAL	TINGGI	TIDAK
7	HP	2020	MURAH	RENDAH	YA
8	HP	2020	MURAH	SEDANG	YA
9	HP	2020	MAHAL	TINGGI	YA
10	DELL	2020	MURAH	RENDAH	YA
11					
486	HP	2023	MURAH	RENDAH	YA
487	HP	2023	MURAH	SEDANG	YA
488	HP	2023	MAHAL	TINGGI	TIDAK
489	DELL	2023	MURAH	RENDAH	TIDAK
490	DELL	2023	MAHAL	SEDANG	TIDAK
491	DELL	2023	MAHAL	TINGGI	TIDAK
492	ACER	2023	MURAH	RENDAH	YA
493	ACER	2023	MAHAL	SEDANG	TIDAK
494	ACER	2023	MAHAL	TINGGI	TIDAK
495	TOSHIBA	2023	MURAH	RENDAH	YA
496	TOSHIBA	2023	MAHAL	SEDANG	TIDAK
497	TOSHIBA	2023	MAHAL	TINGGI	TIDAK
498	ASUS	2023	MURAH	RENDAH	TIDAK
499	ASUS	2023	MAHAL	SEDANG	TIDAK
500	ASUS	2023	MAHAL	TINGGI	TIDAK

Setelah mengimport data, langkah selanjutnya adalah membuat pohon keputusan menggunakan *rapidminer*. Berikut adalah tampilan output dari perhitungan di *rapidminer* dapat dilihat pada gambar 12.



Gambar 12. Tampilan Pohon Keputusan Menggunakan Rapidminer

3.2.11 Menghitung Akurasi Model

Tahap ini memperoleh prediksi untuk data uji menggunakan model yang telah dilatih. Menghitung akurasi model berdasarkan data uji dan menampilkan akurasi model dalam bentuk presentase sehingga dapat dinilai seberapa baik model bekerja dalam memprediksi hasil[12]. Berikut ini adalah tampilan *accuracy model* sebagai berikut :

Result History		ExampleSet (Apply Model)		PerformanceVector (Performance)	
Criterion		Table View		Plot View	
Performance		accuracy		accuracy: 72.00%	
Description					
Annotations					
			true YA	true TIDAK	class precision
		pred. YA	27	18	60.00%
		pred. TIDAK	17	63	78.75%
		class recall	61.36%	77.78%	

Gambar 13. Menghitung Nilai Akurasi

Berdasarkan Gambar 12 diatas menunjukkan bahwa jumlah prediksi untuk true YA-pred.YA sebanyak 27, jumlah prediksi true YA-pred.TIDAK sebanyak 17, jumlah true TIDAK-pred YA sebanyak 18, jumlah true TIDAK-pred.TIDAK sebanyak 63, nilai *class recall* untuk true YA yaitu 61,36%, nilai *class recall* untuk true TIDAK yaitu 77,78% dan nilai *accuracy* dari hasil prediksi adalah 72.00%

3.3 Pembahasan

Implementasi metode Decision Tree untuk prediksi penjualan produk elektronik terlaris di Aries Elektronik dilakukan untuk memprediksi produk-produk terlaris secara otomatis dan sistematis. Tahapan awal yang dilakukan yaitu melakukan pengumpulan data. Data yang digunakan yaitu data penjualan produk elektronik terlaris tahun 2020/2023. Setelah data dikumpulkan tahapan selanjutnya yaitu membagi data menjadi dua yaitu data *training* dan data *testing*. Data *training* yang digunakan

berjumlah 500 data penjualan produk dengan data barang elektronik (merek, jenis barang, tahun, harga kualitas, dan laku). Dalam penelitian ini hanya menggunakan beberapa atribut karena sesuai dengan kriteria yang di gunakan oleh pihak toko dalam memprediksi produk elektronik terlaris. Setelah Data diimplementasikan dengan algoritma C4.5 yang membentuk suatu pohon keputusan(*Decession Tree*) dapat memprediksi produk elektronik yang berpotensi terlaris berdasarkan perhitungan Nilai *Entropy* dan *Information Gain*. Total keseluruhan dari perhitungan nilai *entropy* adalah 0,91861059. Sedangkan nilai *entropy* dan *gain* tertinggi adalah atribut harga(murah) yang dimana nilai *entropy* adalah 0,997597732 dan *information gain* sebanyak 0,088015286 dan data yang diprediksi sebanyak 125 data dengan melakukan *confision matrix* untuk menghitung seberapa tingkat *accuracy*. *Confision matrix* memprediksi suatu data dengan tingkat *accuracy* 72,00% dan memprediksi data sebagai positif(memprediksi produk elektronik paling laris) dan prediksi itu benar. Dalam model berhasil memprediksi 27 produk elektronik paling laris dan model memprediksi data sebagai negatif(memprediksi produk elektronik tidak laris) dan prediksi itu benar. Keputusan yang dihasilkan dari tree ini memberikan panduan yang jelas bagi perusahaan untuk fokus pada produk yang diprediksi akan laris di Toko Aries Elektronik.

4. KESIMPULAN

Berdasarkan hasil penelitian diatas maka dapat diambil kesimpulan bahwa penjualan produk elektronik terlaris diprediksi dengan jumlah data sebanyak 500 data yang diambil dari tahun 2020 sampai 2023. Data yang diambil adalah data penjualan produk elektronik terlaris(laptop) yang dilihat dari beberapa kategori yaitu merek, tahun, harga, kualitas, dan laku. Data ini diimplementasikan dengan algoritma C4.5 yang membentuk suatu pohon keputusan(*Decession Tree*) berdasarkan perhitungan Nilai *Entropy* dan *Information Gain*. Total keseluruhan dari perhitungan nilai *entropy* adalah 0,91861059. Sedangkan nilai *entropy* dan *gain* tertinggi adalah atribut harga(murah) yang dimana nilai *entropy* adalah 0,997597732 dan *information gain* sebanyak 0,088015286 dan data yang diprediksi sebanyak 125 data dengan melakukan *confision matrix* untuk menghitung seberapa tingkat *accuracy*. *Confision matrix* memprediksi suatu data dengan tingkat *accuracy* 72,00% dan memprediksi data sebagai positif(memprediksi produk elektronik paling laris) dan prediksi itu benar. Dalam model berhasil memprediksi 27 produk elektronik paling laris dan model memprediksi data sebagai negatif(memprediksi produk elektronik tidak laris) dan prediksi itu benar. Dan sarannya adalah untuk penelitian yang sama di masa yang akan datang sebaiknya menggunakan algoritma lain, dalam penelitian selanjutnya untuk mencari *accuracy* sebaiknya menggunakan perbandingan metode, dan hasil ini dibuat masih manual, sehingga kedepannya diharapkan dapat dikembangkan ke dalam bentuk aplikasi sistem pengambilan keputusan.

DAFTAR PUSTAKA

- [1] S. P. Dewi and E. Rahayu, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," vol. 3, no. 4, pp. 639–648, 2022, doi: 10.47065/bits.v3i4.1408.
- [2] I. Iddrus and D. W. Sari, "Penerapan Data Mining Menggunakan Algoritma Decision Tree C4.5 Untuk Memprediksi Mahasiswa Drop Out Di Universitas Wiraraja," *J. Adv. Res. Inform.*, vol. 1, no. 02, pp. 1–7, 2023, doi: 10.24929/jars.v1i02.2684.
- [3] Y. Musa and A. Bode, "Prediksi Penjualan Produk Elektronik Menggunakan Metode Least Square," vol. 3, no. 2, pp. 2019–2021, 2019.
- [4] Fairuzabadi Muhammad and Angga Aditya Permana " Sistem Pendukung Keputusan : Konsep, Metode dan Implementasi".Desember 2023. Padang: GET PRESS INDONESIA.2023: 978-623-198-931-4.
- [5] Kelen, Yoseph P.K, *et al.* "Decision support system for the selection of new prospective students using the simple additive weighted (SAW) method." *AIP Conference Proceedings*. Vol. 2798. No. 1. AIP Publishing, 2023
- [6] F. P. Dewanti, S. Setiyowati, and S. Harjanto, "Prediksi Persediaan Obat Untuk Proses Penjualan Menggunakan Metode Decision Tree Pada Apotek," *J. Teknol. Inf. dan Komun.*, vol. 10, no. 1,

- 2022, doi: 10.30646/tikomsin.v10i1.604.
- [7] F. Ilmu *et al.*, “Jurnal ilmiah multidisiplin ilmu,” vol. 1, no. 2, pp. 99–105, 2024.
 - [8] A. E. Wicaksono, “Implementasi Data Mining Dalam Pengelompokan Data Peserta Didik Di Sekolah Untuk Memprediksi Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means (Studi Kasus Sman 16 Bekasi),” *J. Teknol. Rekayasa*, vol. 21, no. 3, pp. 206–216, 2016, [Online]. Available: Jurnal Teknologi Rekayasa
 - [9] D. H. Depari, Y. Widiastiwi, and M. M. Santoni, “Perbandingan Model Decision Tree, Naive Bayes dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung,” *Inform. J. Ilmu Komput.*, vol. 18, no. 3, p. 239, 2022, doi: 10.52958/iftk.v18i3.4694.
 - [10] M. R. Qisthiano, P. A. Prayesy, and I. Ruswita, “Penerapan Algoritma Decision Tree dalam Klasifikasi Data Prediksi Kelulusan Mahasiswa,” *G-Tech J. Teknol. Terap.*, vol. 7, no. 1, pp. 21–28, 2023, doi: 10.33379/gtech.v7i1.1850.
 - [11] S. D. Manullang, E. Buulolo, and I. Lubis, “Implementasi Data Mining Dalam Memprediksi Jumlah Pinjaman Dengan Algoritma C4.5 Pada Kopdit CU Damai Sejahtera,” *J. Sist. Komput. dan Inform.*, vol. 1, no. 3, p. 265, 2020, doi: 10.30865/json.v1i3.2153.
 - [12] M. Sholeh, E. K. Nurnawati, and U. Lestari, “Penerapan Data Mining dengan Metode Regresi Linear untuk Memprediksi Data Nilai Hasil Ujian Menggunakan RapidMiner,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 10–21, 2023, doi: 10.14421/jiska.2023.8.1.10-21.



ZONAsi: Jurnal Sistem Informasi

Is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)