



PENERAPAN DATA MINING UNTUK PREDIKSI PENJUALAN SPANDUK MENGGUNAKAN ALGORITMA C4.5

¹Bagus Triawan*, ²Imran Lubis, ³Lina Arliana Nur Kadim

^{1,2,3}Program Studi Sistem Informasi, Sekolah Tinggi Ilmu Komputer KAPUTAMA
Jl. Veteran No. 4A-9A, Binjai, Sumatera Utara, Indonesia

e-mail :¹ bagustriawan03@gmail.com, ² imran.loebis.medan@gmail.com, ³ lina_arliana@yahoo.com

Received: 2023-08-11

Revised: 2024-02-19

Accepted: 2024-03-08

Page : 149-157

Abstrak : Penjualan spanduk adalah bagian penting dari bisnis periklanan, dan memiliki kemampuan untuk memprediksi penjualan dapat membantu perusahaan dalam perencanaan produksi dan pemasaran yang lebih efektif. Dalam penelitian ini, kami mengumpulkan data penjualan spanduk dari bulan januari hingga bulan juni 2023 dan menggunakan algoritma C4.5 untuk mengolah data tersebut. Dengan metode pohon keputusan dapat membantu dalam menyelesaikan masalah yang terjadi di toko. Rapid miner akan membantu dalam menentukan data barang mana yang lebih laku dan kurang laku. Dengan metode Rapid miner akan mendapatkan data keputusan yang lebih akurat dan mempermudah untuk menganalisa barang. Berdasarkan dari hasil penelitian, spanduk yang sering di pesan oleh konsumen. Hasil dari penelitian ini dapat menjadi panduan bagi perusahaan dalam mengoptimalkan strategi penjualan spanduk mereka.

Kata kunci: Data Mining, Prediksi Penjualan, Algoritma C4.5, Spanduk.

Abstract : Selling banners is an essential part of the advertising business, and having the ability to predict sales can assist companies in more effective production and marketing planning. In this research, we collected banner sales data from January to June 2023 and used the C4.5 algorithm to process the data. The decision tree method can help address issues occurring in the store. RapidMiner will aid in determining which products are more popular and less popular. Using the RapidMiner method will yield more accurate decision data and simplify product analysis. Based on the research findings, banners frequently ordered by consumers. The results of this research can serve as a guide for companies to optimize their banner sales strategies.

Keywords: Data Mining, Sales Prediction, C4.5 Algorithm, Banner.





[Commons Attribution-NonCommercial-ShareAlike 4.0 International License.](#)

1 Pendahuluan

Dalam dunia bisnis, teknologi dapat digunakan sebagai alat untuk memonitor kegiatan-kegiatan yang terjadi dalam bisnis diantaranya transaksi penjualan dan monitoring stok barang dagangan yang tersedia [1]. Penjualan spanduk adalah komponen yang sangat penting dalam industri periklanan, yang sering kali menjadi salah satu elemen utama dalam kampanye promosi produk dan layanan [2]. Kemampuan untuk meramalkan penjualan spanduk dengan akurasi tinggi dapat memberikan keunggulan strategis bagi perusahaan periklanan dalam mengelola inventaris, mengalokasikan sumber daya, dan merencanakan kampanye iklan lebih efektif [3]. Data penjualan historis yang luas dan beragam merupakan sumber daya berharga yang dapat digunakan untuk tujuan ini, tetapi untuk menguraikan pola-pola yang terkandung dalam data tersebut, teknik data mining sangat diperlukan [4]. Data mining adalah pendekatan yang digunakan untuk menggali wawasan dari data historis yang dapat digunakan untuk memprediksi tren penjualan masa depan [5]. Dalam penelitian ini, kami fokus pada penerapan algoritma C4.5 dalam melakukan analisis data mining untuk prediksi penjualan spanduk.

Dalam era digital ini, perusahaan periklanan memiliki akses ke jumlah data yang belum pernah terjadi sebelumnya, yang mencakup informasi seperti perilaku konsumen online, tren sosial media, dan lebih banyak lagi. Namun, penggunaan data semacam ini untuk memprediksi penjualan spanduk tetap menjadi tantangan yang kompleks. Dalam penelitian ini, kami akan memanfaatkan teknik data mining yang canggih, khususnya algoritma C4.5, untuk mengeksplorasi potensi data ini dalam menghasilkan prediksi penjualan spanduk yang akurat dan berguna bagi perusahaan periklanan. Dengan pendekatan ini, kami berharap dapat membantu industri ini mengambil keputusan yang lebih cerdas dan meraih keberhasilan yang lebih besar dalam kampanye periklanan mereka.

2 Tinjauan Literatur (or Literature Review)

Dalam penelitian ini, penulis sedikit banyak terinspirasi dan mereferensi dari penelitian-penelitian sebelumnya yang berkaitan dengan latar belakang masalah pada penelitian ini yaitu:

Menurut penelitian (Kiki Rosita Dewi, et al. 2020) dengan judul “Analisa Algoritma C4.5 untuk Prediksi Penjualan Obat Pertanian di Toko Dewi Sri” menyimpulkan bahwa Penggunaan algoritma C4.5 dalam analisis prediksi penjualan dapat digunakan untuk meramalkan penjualan produk pertanian di masa depan di Toko Dewi Sri. Hal ini didasarkan pada data penjualan yang dikumpulkan dari bulan Oktober hingga November tahun 2019. Dengan penerapan sistem prediksi penjualan menggunakan algoritma C4.5, potensi kesalahan dalam menentukan persediaan obat untuk periode berikutnya dapat diminimalkan. Berdasarkan hasil analisis uji sistem, aplikasi ini memiliki kinerja yang cukup baik dalam melakukan perhitungan peramalan. Dalam kerangka aplikasi prediksi penjualan yang memanfaatkan Algoritma C4.5, tercatat tingkat akurasi sebesar 75% [6].

Dan penelitian yang dilakukan oleh (Aprilianus Kristianus Lalo, et al, 2021) dengan judul ”Implementasi Algoritma C4.5 Untuk Klasifikasi Penjualan Barang di Swalayan Dutalia” menyimpulkan bahwa sistem yang telah dirancang dan dibangun memberikan manfaat yang besar bagi Swalayan Dutalia dengan melakukan klasifikasi menggunakan algoritma C4.5 pada data penjualan masa lalu. Ini membantu dalam pembuatan aturan yang dapat digunakan untuk meramalkan penjualan masa depan, memungkinkan Swalayan untuk menentukan persediaan minimum yang perlu disiapkan. Dengan menggunakan algoritma C4.5, sistem klasifikasi yang telah dirancang dan dibangun berhasil mencapai tingkat akurasi sebesar 100% setelah menguji 50 data hasil prediksi sebagai data uji pada aplikasi RapidMiner Studio Version 9.7. Decision tree yang dihasilkan dalam perhitungan manual, sistem yang dibangun, dan aplikasi RapidMiner adalah sama. Sedangkan penelitian penulis tentang bagaimana memprediksi penjualan spanduk sehingga didapatkan nantinya [7].

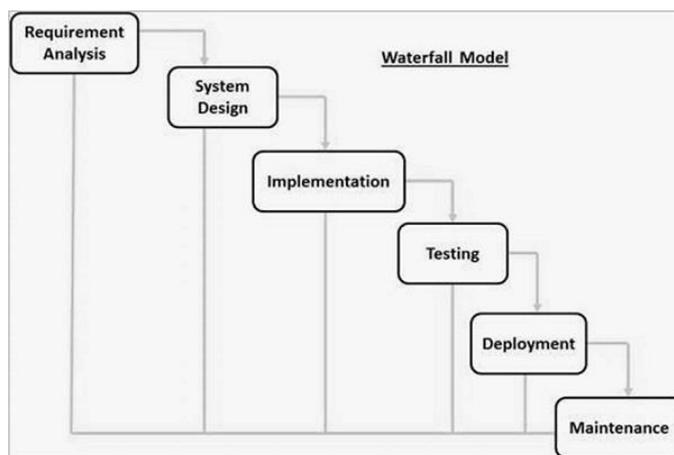
Menurut Arhami & Nasir (2020, h.1) data mining merupakan proses penggalian informasi dan pola yang bermanfaat dari data yang sangat besar. Data mining mencakup pengumpulan data, ekstraksi data, analisis data, dan statistik data. Data *mining* juga dikenal sebagai *knowledge discovery*, *knowledge extraction*, *data/pattern analysis*, *information harvesting*, dan lain-lain [8].

Menurut Sukma, et al (2019, h. 23) Algoritma C4.5 adalah algoritma yang digunakan untuk menghasilkan sebuah pohon keputusan yang dikembangkan oleh Ross quinlan. Ide dasar dari algoritma ini adalah pembuatan pohon keputusan berdasarkan pemilihan atribut yang memiliki prioritas tertinggi atau dapat disebut memiliki nilai gain tertinggi berdasarkan nilai *entropy* atribut tersebut sebagai poros atribut klasifikasi. Pada tahapannya algoritma C4.5 memiliki 2 prinsip kerja, yaitu: Membuat pohon keputusan, dan membuat aturan-aturan (*rule model*). Aturan aturan yang terbentuk dari pohon keputusan akan membentuk suatu kondisi dalam bentuk *if then* [9].

Spanduk merupakan kain rentang berisi slogan, propaganda, atau berita yang perlu diketahui umum. Pengertian lain, spanduk adalah kain membentang yang biasanya berada di tepi jalan, berisi teks, warna, dan gambar mencolok. Spanduk adalah salah satu media informasi, dikarenakan melalui spanduk orang banyak bisa mengetahui terkait promosi tertentu dari suatu tempat usaha, atau mengetahui himbuan terbaru dari pemerintah [10].

3 Metode Penelitian

Dalam penelitian ini, metode yang digunakan adalah metode model waterfall dalam rekayasa perangkat lunak. Model Waterfall adalah suatu pendekatan pengembangan perangkat lunak berurutan di mana proyek dilihat sebagai aliran yang mengalir secara berurutan melalui fase-fase perencanaan, pemodelan, implementasi (konstruksi), dan pengujian, serupa dengan aliran air dalam air terjun [11]. Tahap-tahap dalam model penelitian *waterfall* adalah sebagai berikut :



Gambar 1 Tahap-Tahap Metode *Waterfall*

Pengambilan data dilakukan setelah penulis mendapatkan ijin dari pemilik Toko Dian Photo Advertising Binjai. Adapun data yang didapatkan transaksi penjualan dari bulan januari hingga juni 2023. Adapun data tersebut yaitu seperti pada tabel dibawah ini.

Tabel 1. Data Penjualan Spanduk

No	Bulan	Jenis Spanduk	Bahan	Harga	Jumlah
1	Januari	Outdoor printing	Fleksi German	Rp 28.000	15
2	Januari	Outdoor printing	Fleksi Korea	Rp 16.000	13
3	Januari	Outdoor printing	Fleksi Cina	Rp 13.000	24
4	Januari	Outdoor printing	Albatros	Rp 25.000	21
5	Januari	Outdoor printing	Luster	Rp 25.000	39
6	Januari	Outdoor printing	Easy Banner	Rp 26.000	22

7	Januari	Outdoor printing	Poly Clouth	Rp 25.000	26
8	Januari	Indoor printing	Fleksi German	Rp 132.000	7
9	Januari	Indoor printing	Fleksi Korea	Rp 99.000	2
10		Indoor printing	Fleksi Cina	Rp 77.000	17
...
126	Juni	UV Printing	Poly Clouth	Rp. 154.000	9

Proses transformasi data dilakukan untuk mengubah atribut sesuai dengan format yang dapat diproses dalam program. Atribut jenis spanduk, bahan spanduk dan harga dilakukan perubahan nilai atribut dari diskrit menjadi numerik (label) pada program sehingga didapatkan label numerik pada Tabel 2 berikut.

Tabel 2 Transformasi Data

Variabel	Atribut	Transformasi Data
Jenis spanduk	Indoor printing	0
	Outdoor printing	1
	UV Printing	2
Bahan	Fleksi German	0
	Fleksi Korea	1
	Fleksi Cina	2
	Albatros	3
	Luster	4
	Easy Banner	5
	Poly Clouth	6
Harga	<= Rp.75.000	Murah
	Rp.76.000-Rp150.000	Normal
	>= Rp 151.000	Mahal

Selain atribut, kelas / label *output* yang digunakan juga dilakukan perubahan nilai numerik ke nilai diskrit dengan mencari nilai rata-rata dari total penjualan pada data tersebut.

$$\text{rata-rata} = \frac{\text{total terjual}}{\text{banyak data}}$$

$$\text{rata-rata} = \frac{1501}{126} = 11,9 = 12$$

Setelah didapatkan nilai rata-rata dari data tersebut, ketentuan untuk perubahan nilai numerik ke diskrit adalah sebagai berikut :

Terjual ≤ 12 pcs : tidak laris

Terjual > 12 pcs : laris

Tabel 3 adalah data yang telah dilakukan perubahan nilai tabel kelas numerik menjadi nilai label kelas diskrit dan menjadi data *training*. Data *training* dan data prediksi yang di-*input* harus mencakup atribut dan nilai atribut tersebut.

Tabel 3 Data Training

No	Bulan	Jenis Spanduk	Bahan	Harga	Jumlah	Status Penjualan
1	Januari	0	0	Murah	15	Laris
2	Januari	0	1	Murah	13	Laris
3	Januari	0	2	Murah	24	Laris

4	Januari	0	3	Murah	21	Laris
5	Januari	0	4	Murah	39	Laris
6	Januari	0	5	Murah	22	Laris
7	Januari	0	6	Murah	26	Laris
8		1	0	Normal	7	Tidak Laris
9		1	1	Normal	2	Tidak Laris
10		1	2	Normal	17	Laris
...
126	Juni	2	6	Mahal	9	Tidak Laris

Data training digunakan dalam proses *training* untuk menemukan pola penjualan menggunakan Algoritma C4.5. Pola yang dihasilkan berupa pohon keputusan yang berisikan *rules* penjualan. *Rules* yang dihasilkan akan digunakan untuk melakukan prediksi status penjualan barang yang baru.

Tabel 4 Data Testing

No	Jenis Spanduk	Bahan	Harga	Jumlah
1	Outdoor printing	Fleksi Korea	<= Rp.75.000	66
2	Outdoor printing	Fleksi Cina	<= Rp.75.000	83
3	Indoor printing	Luster	Rp.76.000-Rp150.000	75
4	Indoor printing	Easy Banner	Rp.76.000-Rp150.000	33
5	UV printing	Poly Clouth	>= Rp 150.100	46

Data prediksi berisikan data barang beserta atribut yang memengaruhi omzet penjualannya. Data prediksi belum memiliki status penjualan, sehingga harus dilakukan prediksi untuk menentukan status penjualannya.

4 Hasil dan Pembahasan (or Results and Analysis)

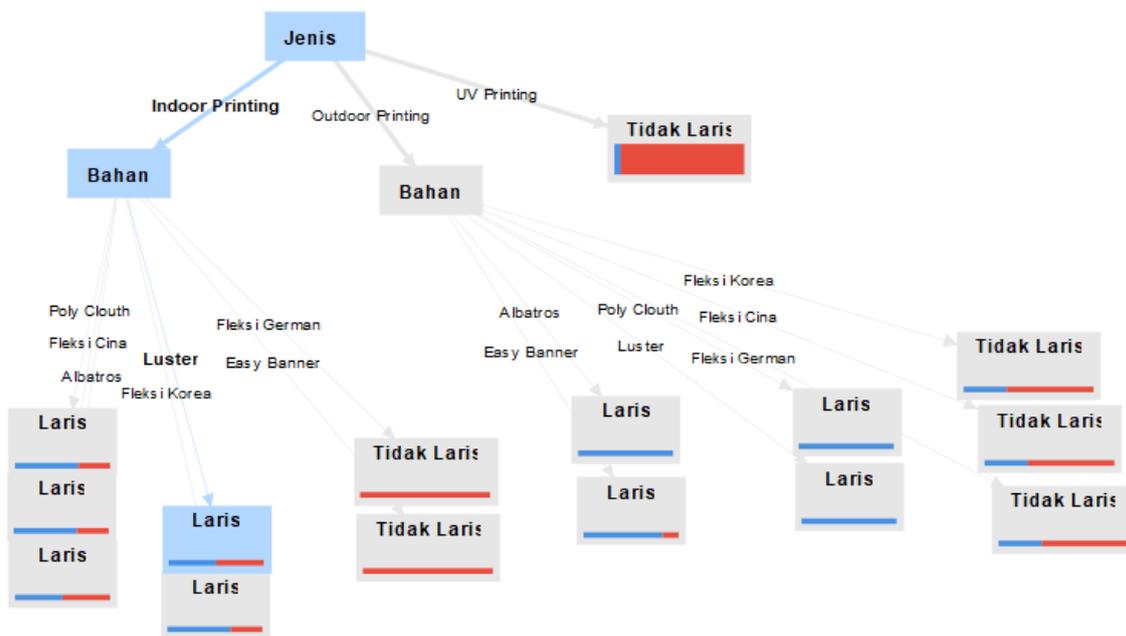
4.1 Perhitungan Root Decision Tree Dari Data Training

Tabel 5 Data Sesuai Target Pola

Node	Atribut	Nilai	Jumlah kasus	Laris	Tidak Laris	entropy	gain
	Total		126	49	77	0,9641	
	Jenis Spanduk	Outdoor	42	29	13	0.89262	
		Indoor	42	18	24	0.98523	
		UV	42	2	40	0.2762	
							0.24606
	Jenis Bahan	Fleksi German	18	2	16	0,5033	
		Fleksi Korea	18	7	11	0,9641	
		Fleksi Cina	18	6	12	0,9183	
		Albatros	18	9	9	1,0000	
		Luster	18	10	8	0,9911	
		Easy Banner	18	5	13	0,8524	
		Poly Clouth	18	10	8	0,9911	
							0.07548

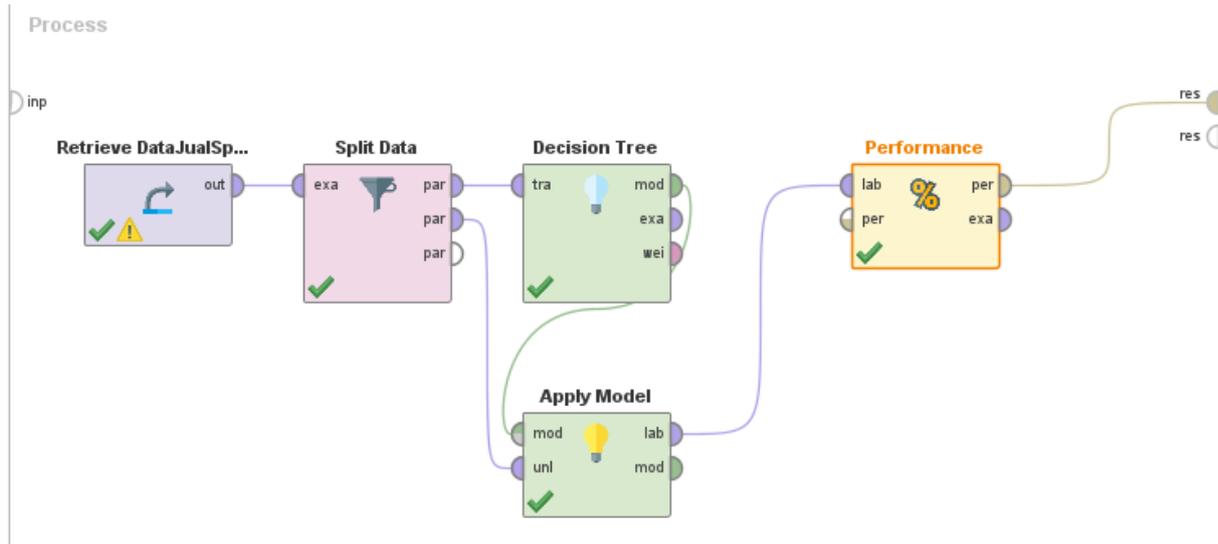
	Harga	Murah	42	29	13	0,8926	
		Normal	54	19	35	0,9357	
		Mahal	30	1	29	0,2108	
							0.2153

Dari hasil perhitungan nilai *gain* didapatkan hasil bahwa nilai *gain* jenis spanduk lebih besar dibanding atribut lainnya sehingga atribut jenis spanduk dijadikan node akar dan mempunyai 3 cabang yaitu, Indoor printing, outdoor printing, UV printing. Berikut adalah model *decision tree* dengan menggunakan data set penjualan spanduk selama 6 bulan.



Gambar 2 Model *Decision Tree*

Langkah selanjutnya adalah melakukan proses pengujian terhadap data test melalui tools Rapidminer 5 yang dimulai dari proses koneksi antara basis data, operator, dan validasi seperti gambar 3 berikut ini.



Gambar 3 Koneksi Proses Rapidminer 5

Dari proses koneksi pada gambar 3 di atas didapatkan pohon keputusan yang sama dengan gambar 2 dan hasil pengukuran akurasi. Berikut adalah screenshot hasil pengukuran akurasi terhadap kinerja algoritma C.45 dalam memprediksi penjualan spanduk yaitu sebesar 84%.

Table View
 Plot View

accuracy: 84.00%

	true Laris	true Tidak Laris	class precision
pred. Laris	7	1	87.50%
pred. Tidak Laris	3	14	82.35%
class recall	70.00%	93.33%	

Gambar 4 Screenshot Akurasi Kinerja Algoritma C.45

Selanjutnya setelah proses ini dilakukan penarikan informasi yang akan menjadi pengetahuan baru berdasarkan pohon keputusan pada Rapidminer 5.

Tree

```
Jenis Spanduk = Indoor
| Bahan = Albatros: Laris {Tidak Laris=0, Laris=1}
| Bahan = Easy Banner: Tidak Laris {Tidak Laris=1, Laris=0}
| Bahan = Fleksi Cina: Laris {Tidak Laris=0, Laris=1}
| Bahan = Fleksi German: Tidak Laris {Tidak Laris=1, Laris=0}
| Bahan = Fleksi Korea: Laris {Tidak Laris=0, Laris=1}
| Bahan = Luster: Laris {Tidak Laris=0, Laris=1}
| Bahan = Poly Clouth: Laris {Tidak Laris=0, Laris=1}
Jenis Spanduk = Outdoor
| Bahan = Albatros: Laris {Tidak Laris=0, Laris=1}
| Bahan = Easy Banner: Laris {Tidak Laris=0, Laris=1}
| Bahan = Fleksi Cina: Laris {Tidak Laris=0, Laris=1}
| Bahan = Fleksi German: Tidak Laris {Tidak Laris=1, Laris=0}
| Bahan = Fleksi Korea: Tidak Laris {Tidak Laris=1, Laris=0}
| Bahan = Luster: Laris {Tidak Laris=0, Laris=1}
| Bahan = Poly Clouth: Laris {Tidak Laris=0, Laris=1}
Jenis Spanduk = UV: Tidak Laris {Tidak Laris=7, Laris=0}
```

Gambar 5. Pengetahuan baru yang dihasilkan melalui Rapidminer 5

5 Kesimpulan

Melalui proses data mining yang telah dilakukan menggunakan Algoritma C.45 dengan aplikasi *Rapidminer 5* diketahui informasi sebagai pengetahuan baru berupa 6 *rule*/aturan dengan tingkat kinerja algoritma C.45 sebesar 84%. *Rule* yang ditemukan yang dapat dijadikan landasan bagi manajemen perusahaan dalam melakukan prediksi pencapaian target penjualan sehingga pengelola perusahaan dapat mengantisipasi dengan mengambil langkah - langkah bisnis yang tepat dalam mengupayakan tercapainya penjualan dan penyediaan bahan baku.

Referensi

- [1] A. Cahyo, "Perancangan Aplikasi Inventory Control PT. Kimia Farma Apotek," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 4, no. 3, pp. 288–296, 2020.



- [2] N. Anisa Fitri, “Dampak E-Commerce Terhadap Strategi Pemasaran (Studi Kasus Pada Platform Shopee),” *PROCURATIO : Jurnal Manajemen & Bisnis*, vol. 01, no. 2, pp. 67–77, 2022.
- [3] S. Aman, H. Fitria, and C. B. Hakim, “Analisa Peramalan Penjualan Tas Kerajinan Ekspor Dengan Metode Moving Average Di CV. Sindocraft,” *Jurnal Teknologi dan Manajemen Industri*, vol. 4, no. 1, pp. 1–6, 2023.
- [4] R. A. SUkomoni and Supardi, *Buku Ajar Manajemen Operasional Dan Implementasi Dalam Industri*, vol. I. Sidoarjo: UMSIDA Press, 2020.
- [5] K. Hamidah and A. Voutama, “Analisis Faktor Tingkat Kebahagiaan Negara Menggunakan Data World Happiness Report dengan Metode Regresi Linier,” *Explore IT*, vol. 15, no. 1, pp. 1–7, 2023, doi: 10.35891/explorit.
- [6] K. R. Dewi, K. F. Mauladi, and Masruroh, “Analisa Algoritma C4.5 untuk Prediksi Penjualan Obat Pertanian di Toko Dewi Sri,” in *Seminar Nasional Inovasi Teknologi*, Kediri: UN PGRI Kediri, Jul. 2020, pp. 109–114.
- [7] A. K. Lalo, P. Batarius, and Y. C. H. Siki, “Implementasi Algoritma C4.5 Untuk Klasifikasi Penjualan Barang di Swalayan Dutalia,” *Jurnal Teknik Informatika Unika St. Thomas (JTIUST)*, vol. 6, no. 1, pp. 1–12, 2021.
- [8] M. Arhami and M. Nasir, *Data Mining Algoritma dan Implementasi*. Yogyakarta: CV Andi Offset, 2020.
- [9] A. R. Sukma, R. Halfis, and A. Hermawan, “Klasifikasi Channel Youtube Indonesia Menggunakan Algoritma C4.5,” *Jurnal Teknik Komputer AMIK BSI*, vol. 5, no. 1, pp. 21–28, 2019.
- [10] N. N. Rahman, R. S. Utami, and A. Z. M. Raden, “Analisis Tipografi Pada Spanduk Pemasaran Perumahan,” *Jurnal Desain*, vol. 5, no. 3, pp. 250–257, 2018.
- [11] C. Trisianto, “Penggunaan Metode Waterfall Untuk Pengembangan Sistem Monitoring dan Evaluasi Pembangunan Pedesaan,” *Jurnal Teknologi Informasi ESIT*, vol. XII, no. 1, pp. 8–22, 2018.