

KLASIFIKASI FAKTOR PENYEBAB PERCERAIAN MENGUNAKAN ALGORITMA C4.5

Agnes Irene Silitonga¹, Darwita Arini Nasution², Histi Trifesi Naibaho³,
Selvia Apriani⁴, Dewi Pika Lumban Bantu⁵

^{1,2,3,4}Program Studi Bisnis Digital, Fakultas Ekonomi, Universitas Negeri Medan

¹Program Studi Pendidikan Pancasila dan Kewarganegaraan, Fakultas Ilmu Sosial, Universitas Negeri Medan

¹agnesirenesilitonga@unimed.ac.id, ²darwitaarini@gmail.com, ³histinaibaho@gmail.com, ⁴selviaapriani456@gmail.com,

⁵dewi_pika_lumban@unimed.ac.id

ABSTRACT

Divorce is a complex social phenomenon that has significant impacts on individuals and society. In Indonesia, divorce is a serious concern for the government and society because the divorce rate has continued to increase in recent decades. Understanding the factors that contribute to divorce is essential to developing effective prevention strategies. The objective of this study is to classify the factors of divorce in Indonesia using the C4.5 algorithm, an algorithm used for data mining in building decision trees. This study used divorce data from 20 provinces in Indonesia in 2023 with various factors causing divorce. Divorce data was obtained from the Central Statistics Agency of Indonesia. The research process includes data collection, data pre-processing, and application of the C4.5 algorithm to build a classification model. The results of the study showed that continuous disputes and quarrels, apostasy, and economy are the most significant factors in divorce. The research results are expected to be a reference for policy makers and marriage counselors in formulating appropriate interventions to reduce divorce rates. The findings of this study can also serve as a foundation for establishing more effective policies and interventions to reduce divorce rates and enhance family institutions in Indonesia.

Keywords- C4.5 Algorithm, Data Mining, Decision Tree, Divorce Factors

I. PENDAHULUAN

Fenomena perceraian merupakan masalah sosial yang kompleks dan terus mengalami peningkatan di Indonesia. Berdasarkan data yang diperoleh dari Badan Pusat Statistik (BPS), dalam beberapa dekade terakhir, jumlah perceraian terus meningkat di Indonesia yang mana hal ini memicu berbagai kajian untuk memahami faktor-faktor penyebab dan dampaknya. Laporan Statistik Indonesia Tahun 2024 yang dirilis oleh Badan Pusat Statistik, data menunjukkan bahwa pada tahun 2023, terdapat lebih dari 400.000 kasus perceraian yang tercatat di seluruh Indonesia, yang sebagian besar disebabkan oleh faktor ekonomi, perselingkuhan, dan ketidakcocokan antar pasangan [1].

Selain berdampak terhadap pasangan yang bercerai, fenomena perceraian juga berdampak pada anak-anak, keluarga besar, dan lingkungan sosial mereka. Perceraian dapat memicu permasalahan emosional pada anak, menurunkan kesejahteraan keluarga, serta menimbulkan beban sosial dan ekonomi yang tinggi bagi masyarakat [2]. Sangat penting untuk memahami faktor-faktor ini secara mendalam guna merumuskan strategi pencegahan dan intervensi yang efektif. Namun, kompleksitas dan interaksi antar faktor membuat analisis secara manual menjadi sulit dan kurang efisien.

Dalam konteks inilah teknologi data mining menjadi relevan. Data mining adalah proses ekstraksi informasi penting dari kumpulan data besar melalui metode statistik, matematika, dan pembelajaran mesin [3]. Salah satu algoritma data mining yang efektif untuk klasifikasi adalah algoritma C4.5. Algoritma ini digunakan untuk membangun pohon keputusan yang

dapat membantu mengidentifikasi pola dan hubungan antar variabel dalam dataset [4].

Algoritma C4.5 dikembangkan oleh Quinlan sebagai pengembangan dari algoritma ID3 dengan kemampuan menangani atribut kontinu dan menangani data yang hilang [5]. Algoritma ini mampu menangani data dengan atribut yang kontinu dan diskrit, serta menghasilkan model yang mudah diinterpretasikan. Dalam penelitian sebelumnya, algoritma C4.5 telah berhasil diterapkan dalam berbagai bidang seperti prediksi penyakit [6], klasifikasi gizi buruk dan *stunting* [7], klasifikasi kepuasan pelanggan [8], analisis kredit [9], dan penentuan penerima beasiswa [10]. Hal ini menunjukkan potensi algoritma C4.5 dalam mengolah data yang kompleks dan menghasilkan model yang akurat.

Dalam konteks penelitian ini, algoritma C4.5 digunakan untuk mengklasifikasikan faktor-faktor utama penyebab perceraian di Indonesia. Penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam literatur mengenai perceraian di Indonesia. Temuan dari penelitian ini tidak hanya akan memperkaya pemahaman akademis tentang faktor-faktor penyebab perceraian, tetapi juga dapat menjadi dasar bagi pengembangan kebijakan publik yang lebih efektif. Dengan demikian, penelitian ini juga diharapkan dapat membantu menurunkan angka perceraian dan meminimalkan dampak negatifnya terhadap masyarakat.

II. LANDASAN TEORI

A. Data Mining

Data mining adalah proses menyortir sebuah set big data untuk mengidentifikasi pola dan hubungan yang

dapat membantu memecahkan masalah melalui analisis data [11]. Teknik dan *tools* dalam *data mining* digunakan untuk memprediksi tren masa depan dan membuat keputusan yang lebih tepat.

Sebagai salah satu disiplin ilmu data, *data mining* menggunakan teknik analisis tingkat lanjut untuk menemukan informasi yang berguna dalam set data [12]. Data mining adalah langkah penting dalam proses yang lebih luas yang dikenal sebagai *Knowledge Discovery in Databases* (KDD) yang mencakup serangkaian langkah mulai dari pembersihan data hingga interpretasi pola yang ditemukan [13].

B. Decision Tree

Decision tree merupakan struktur partisi rekursif yang digunakan dalam data mining untuk memecahkan masalah klasifikasi dan regresi dengan representasi berbentuk pohon [14]. Setiap cabang pohon ini melambangkan aturan keputusan yang membawa pengguna dari fitur ke fitur hingga mencapai kesimpulan berupa hasil klasifikasi atau prediksi. Karena sifatnya yang mudah dipahami, metode ini sering digunakan dalam pengambilan keputusan yang melibatkan data yang kompleks atau data yang sulit dipahami secara langsung.

Pada *decision tree*, data ditransformasikan menjadi pohon keputusan, yang kemudian diubah menjadi aturan. *Decision tree* menyimpan data dalam bentuk tabel yang berisi *record* dan atribut yang mana atribut merupakan parameter yang digunakan sebagai standar pembentukan *tree*. *Root* yang tidak memiliki cabang masukan memiliki dampak paling besar pada suatu kelas sedangkan *leaves* merupakan sambungan *node* dari sebuah *tree* yang memiliki label kelas, dan simpul internal menunjukkan test atau subset atribut.

C. Algoritma C4.5

Algoritma C4.5 adalah salah satu algoritma pembelajaran mesin yang dikembangkan oleh Ross Quinlan pada tahun 1993 dan digunakan untuk membentuk pohon keputusan, yang secara luas diaplikasikan dalam masalah klasifikasi. Algoritma ini menghasilkan pohon keputusan yang mudah dipahami karena menyajikan proses pengambilan keputusan dalam bentuk percabangan logis. Oleh karena itu, algoritma ini banyak digunakan dalam data mining, penggalian informasi, serta pengambilan keputusan yang melibatkan data besar atau kompleks.

Algoritma C4.5 bekerja berdasarkan informasi yang terdapat dalam atribut atau fitur yang tersedia pada dataset untuk membentuk pohon keputusan. Pohon keputusan ini digunakan untuk mengklasifikasikan data baru ke dalam kategori yang telah ditentukan. Setiap simpul di dalam pohon akan mewakili atribut atau fitur dari dataset, sementara setiap cabang mewakili nilai yang mungkin dari atribut tersebut. Proses ini terus berlanjut hingga mencapai simpul daun, yang merepresentasikan kelas atau label akhir yang dituju.

Penentuan atribut akar ditentukan oleh atribut dengan nilai *gain* tertinggi. Untuk menghitung *gain* digunakan persamaan (1).

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

dimana A adalah atribut yang dipertimbangkan, $Values(A)$ adalah nilai-nilai yang mungkin dari atribut A , S_v adalah subset dari S yang memiliki nilai $A=v$, S_v adalah jumlah data dalam S_v , dan $|S|$ adalah jumlah total data dalam S .

Sementara itu, untuk menghitung nilai *Entropy* digunakan persamaan (2).

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

dimana A adalah atribut yang dipertimbangkan, $Values(A)$ adalah nilai-nilai yang mungkin dari atribut A , S_v adalah subset dari S yang memiliki nilai $A=v$, S_v adalah jumlah data dalam S_v , dan $|S|$ adalah jumlah total data dalam S .

III. METODOLOGI PENELITIAN

Penelitian ini menggunakan algoritma C4.5 untuk mengklasifikasikan faktor-faktor yang mempengaruhi perceraian. Pada penelitian ini, digunakan data perceraian yang diperoleh dari Badan Pusat Statistik Tahun 2023 [15]. Variabel-variabel yang dianalisis meliputi faktor penyebab perceraian seperti zina, madat (narkoba), judi, poligami, KDRT, Perselisihan dan pertengkarannya, teruserus, murtad (keluar dari agama Islam) dan ekonomi. Tahapan penelitian ini meliputi:

1. Pengumpulan data
Data perceraian dikumpulkan melalui dataset yang tersedia di Badan Pusat Statistik (BPS).
2. *Pre-processing* data
Data yang dikumpulkan kemudian diproses untuk ditransformasikan atau diubah ke dalam bentuk yang sesuai, agar dapat diproses dengan algoritma C4.5. Langkah ini penting untuk memastikan kualitas dan keakuratan data yang digunakan dalam analisis.
3. Perhitungan *entropy* dan informasi *gain*
Perhitungan *entropy* dilakukan untuk menghitung rasio perolehan informasi yang ada dalam data. Perhitungan *gain* dilakukan untuk mengetahui atribut dengan informasi *gain* tertinggi dan menentukan atribut mana yang akan dijadikan simpul akar pada *decision tree*.
4. Pohon keputusan (*Decision tree*)
Setelah *entropy* dan *gain* information diperoleh, proses pembentukan pohon keputusan dilakukan sampai semua atribut pada pohon keputusan memiliki kelas dan proses perhitungan tidak dapat lagi dilakukan.
5. Aturan / *rules*
Pohon keputusan yang dihasilkan direpresentasikan dalam bentuk uraian penjelasan atau disebut juga aturan / *rules*.
6. Interpretasi hasil
Hasil analisis pohon keputusan diinterpretasikan untuk mengidentifikasi faktor-faktor utama yang berkontribusi terhadap perceraian. Atribut yang memiliki pengaruh signifikan akan dibahas lebih lanjut untuk memberikan wawasan yang lebih mendalam.

A. Analisa Data

Data yang digunakan adalah data perceraian dari 20 provinsi beserta faktor penyebab perceraian seperti pada Tabel 1. Kemudian dilakukan *pre-processing* pada data perceraian berdasarkan sembilan atribut yang mana atribut ini merupakan faktor penyebab perceraian diantaranya perzinahan, madat (narkoba), judi, poligami,

KDRT, perselisihan dan pertengkaran terus menerus, murtad (keluar dari Agama Islam), ekonomi, serta jumlah.

Tabel 1. Data Perceraian

Provinsi	Zina	Madat	Judi	Poligami	Kekerasan Dalam Rumah Tangga	Perselisihan dan Pertengkaran Terus Menerus	Murtad	Ekonomi	Jumlah
Aceh	0	12	25	25	102	4915	12	235	5326
Sumatera Utara	1	15	121	8	136	13709	73	465	14528
Sumatera Barat	4	2	3	10	19	6969	18	113	7138
Riau	7	25	23	9	71	7333	23	231	7722
Sumatera Selatan	5	35	48	30	259	8311	50	574	9312
Lampung	8	8	81	15	167	10099	47	2838	13263
DKI Jakarta	14	25	57	43	178	9464	92	2452	12325
Jawa Tengah	39	20	143	40	235	36618	174	23176	60445
DI Yogyakarta	4	3	7	7	58	3921	21	583	4604
Jawa Timur	579	38	415	132	1636	35940	156	33572	72468
Bali	3	1	2	1	15	886	26	34	968
Nusa Tenggara Barat	4	5	15	29	240	4997	23	198	5511
Nusa Tenggara Timur	3	0	3	0	19	320	3	8	356
Kalimantan Barat	6	18	33	28	86	3750	23	312	4256
Kalimantan Tengah	2	5	9	18	35	2310	11	178	2568
Sulawesi Selatan	22	28	60	61	333	9856	33	399	10792
Gorontalo	7	3	1	3	62	1452	9	40	1577
Sulawesi Barat	4	2	5	1	39	712	0	33	796
Maluku	0	1	0	4	29	632	1	8	675
Papua	0	1	7	2	27	889	14	32	972

Tabel 2. Atribut Perzinahan

Perzinahan	<i>Pre-processing</i>
0-100	Z1
101-200	Z2
201-300	Z3
301-400	Z4
401-500	Z5
501-600	Z6

Tabel 3. Atribut Madat (Narkoba)

Madat (Narkoba)	<i>Pre-processing</i>
0-5	M1
6-10	M2
11-15	M3
16-20	M4
21-25	M5
26-30	M6
31-35	M7
36-40	M8

Tabel 4. Atribut Judi

Judi	<i>Pre-processing</i>
0-100	J1
101-200	J2
201-300	J3
301-400	J4
401-500	J5

Tabel 5. Atribut Poligami

Poligami	<i>Pre-processing</i>
0-25	P1
26-50	P2
51-75	P3
76-100	P4
101-125	P5
126-150	P6

Tabel 6. Atribut KDRT

Kekerasan Dalam Rumah Tangga (KDRT)	<i>Pre-processing</i>
0-200	KDRT1
201-400	KDRT2
401-600	KDRT3
601-800	KDRT4
801-1000	KDRT5
1001-1200	KDRT6
1201-1400	KDRT7

1401-1600	KDRT8
1601-1800	KDRT9
1801-2000	KDRT10

Tabel 7. Atribut PPTM

Perselisihan dan Pertengkaran Terus Menerus (PPTM)	<i>Pre-processing</i>
0-5000	PPTM1
5001-10000	PPTM2
10001-15000	PPTM3
15001-20000	PPTM3
20001-25000	PPTM4
25000-30000	PPTM5
30001-35000	PPTM6
35001-40000	PPTM7

Tabel 8. Atribut Murtaf

Murtaf	<i>Pre-processing</i>
0-20	MM1
21-40	MM2
41-60	MM3
61-80	MM4
81-100	MM5
101-120	MM6
121-140	MM7
141-160	MM8
161-180	MM9
181-200	MM10

Tabel 9. Atribut Ekonomi

Ekonomi	<i>Pre-processing</i>
0-5000	E1
5001-10000	E2
10001-15000	E3
15001-20000	E4
20001-25000	E5
25000-30000	E6
30001-35000	E7

Tabel 10. Atribut Jumlah Perceraian

Jumlah	<i>Pre-processing</i>
0-40000	Normal
40001-80000	Tidak Normal

Setelah dilakukan *pre-processing*, selanjutnya dilakukan transformasi data berdasarkan atributnya masing – masing seperti pada Tabel 11.

Tabel 3. Data setelah dilakukan *pre-processing*

Provinsi	Zina	Madat	Judi	Poligami	Kekerasan Dalam Rumah Tangga (KDRT)	Perselisihan dan Pertengkaran Terus Menerus (PPTM)	Murtad	Ekonomi	Jumlah
Aceh	Z1	M3	J1	P1	KDRT1	PPTM1	MM1	E1	NORMAL
Sumatera Utara	Z1	M3	J2	P1	KDRT1	PPTM3	MM4	E1	NORMAL
Sumatera Barat	Z1	M1	J1	P1	KDRT1	PPTM2	MM1	E1	NORMAL
Riau	Z1	M5	J1	P1	KDRT1	PPTM2	MM2	E1	NORMAL
Sumatera Selatan	Z1	M7	J1	P2	KDRT2	PPTM2	MM3	E1	NORMAL
Lampung	Z1	M2	J1	P1	KDRT1	PPTM3	MM3	E1	NORMAL
DKI Jakarta	Z1	M5	J1	P2	KDRT1	PPTM2	MM5	E1	NORMAL
Jawa Tengah	Z1	M4	J2	P2	KDRT2	PPTM7	MM9	E5	TIDAK NORMAL
DI Yogyakarta	Z1	M1	J1	P1	KDRT1	PPTM1	MM2	E1	NORMAL
Jawa Timur	Z6	M8	J5	P6	KDRT9	PPTM7	MM8	E7	TIDAK NORMAL
Bali	Z1	M1	J1	P1	KDRT1	PPTM1	MM2	E1	NORMAL
Nusa Tenggara Barat	Z1	M1	J1	P2	KDRT2	PPTM1	MM2	E1	NORMAL
Nusa Tenggara Timur	Z1	M1	J1	P1	KDRT1	PPTM1	MM1	E1	NORMAL
Kalimantan Barat	Z1	M4	J1	P2	KDRT1	PPTM1	MM2	E1	NORMAL
Kalimantan Tengah	Z1	M1	J1	P1	KDRT1	PPTM1	MM1	E1	NORMAL
Sulawesi Selatan	Z1	M6	J1	P3	KDRT2	PPTM2	MM2	E1	NORMAL
Gorontalo	Z1	M1	J1	P1	KDRT1	PPTM1	MM1	E1	NORMAL
Sulawesi Barat	Z1	M1	J1	P1	KDRT1	PPTM1	MM1	E1	NORMAL
Maluku	Z1	M1	J1	P1	KDRT1	PPTM1	MM1	E1	NORMAL
Papua	Z1	M1	J1	P1	KDRT1	PPTM1	MM1	E1	NORMAL

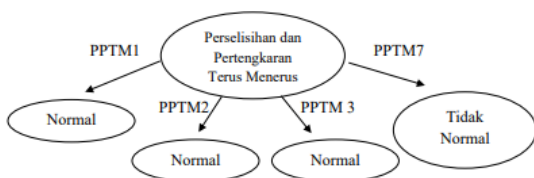
Tabel 11 merupakan hasil transformasi data, dari 20 provinsi diperoleh 18 provinsi dikategorikan normal dan 2 provinsi dikategorikan tidak normal. Selanjutnya dilakukan perhitungan *entropy* dan *gain* untuk setiap atribut seperti pada Tabel 12.

Tabel 4. Hasil *Entropy* dan *Gain*

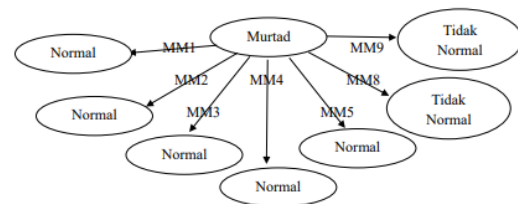
Atribut	Keterangan	Total Kasus	Normal	Tidak Normal	Entropy	Gain
Total		20	18	2	0.47	
Zina						0.19
	Z1	19	18	1	0.29	
	Z6	1	0	1	0	
Madat						0.37
	M1	10	10	0	0	
	M2	1	1	0	0	
	M3	2	2	0	0	

	M4	2	1	1	1	
	M5	2	2	0	0	
	M6	1	1	0	0	
	M7	1	1	0	0	
	M8	1	0	1	0	
Judi						0.37
	J1	17	17	0	0	
	J2	2	1	1	1	
	J5	1	0	1	0	
Poligami						0.09
	P1	13	13	10	0.29	
	P2	5	4	1	0.72	
	P3	1	1	0	0	
	P6	1	0	1	0	
Kekerasan Dalam Rumah Tangga						0.31
	KDRT1	15	15	0	0	
	KDRT2	4	3	1	0.81	
	KDRT9	1	0	1	0	
Perselisihan dan Pertengkaran Terus Menerus						0.47
	PPTM1	11	11	0	0	
	PPTM2	5	5	0	0	
	PPTM3	2	2	0	0	
	PPTM7	2	0	2	0	
Murtad						0.47
	MM1	8	8	0	0	
	MM2	6	6	0	0	
	MM3	2	2	0	0	
	MM4	1	1	0	0	
	MM5	1	1	0	0	
	MM8	1	0	1	0	
	MM9	1	0	1	0	
Ekonomi						0.47
	E1	18	18	0	0	
	E5	1	0	1	0	
	E7	1	0	1	0	

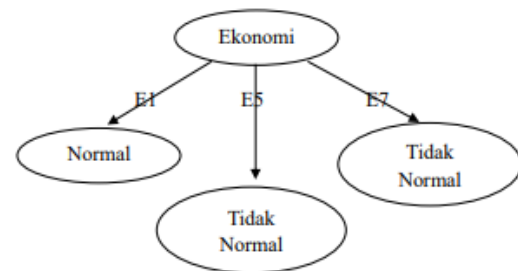
Berdasarkan tabel 12 diperoleh 3 atribut yang memiliki nilai *gain* tertinggi dengan nilai 0,4689955936 yaitu perselisihan dan pertengkaran terus menerus (PPTM), murtad (keluar dari agama Islam) dan ekonomi. Dengan demikian ketiga atribut dapat menjadi node akar. Maka dapat dibentuk pohon keputusan seperti pada gambar 1, 2, dan 3.



Gambar 1. Pohon keputusan node akar PPTM



Gambar 2. Pohon keputusan node akar murtad



Gambar 3. Pohon keputusan node akar ekonomi

Proses pembentukan pohon keputusan berhenti karena tidak ada lagi atribut *root* / akar yang akan terbentuk berdasarkan hasil perhitungan *entropy* dan *gain*. Gambar 3 merupakan pohon keputusan terakhir yang terbentuk dan diperoleh hasil klasifikasi dalam bentuk *rules*.

Pada gambar 1, diperoleh 4 aturan / *rules* yang terbentuk dari perhitungan *gain* dan *entropy*, diantaranya:

1. PPTM dengan PPTM1 dianggap normal
2. PPTM dengan PPTM2 dianggap normal
3. PPTM dengan PPTM3 dianggap normal
4. PPTM dengan PPTM7 dianggap tidak normal

Berdasarkan gambar 2, diperoleh 7 aturan / *rules* yang terbentuk dari perhitungan *gain* dan *entropy* diantaranya

1. Murtad dengan MM1 dianggap normal
2. Murtad dengan MM2 dianggap normal
3. Murtad dengan MM3 dianggap normal
4. Murtad dengan MM4 dianggap normal
5. Murtad dengan MM5 dianggap normal
6. Murtad dengan MM8 dianggap tidak normal
7. Murtad dengan MM9 dianggap tidak normal

Berdasarkan gambar 3, diperoleh tiga aturan / *rules* yang terbentuk dari perhitungan *gain* dan *entropy* yaitu:

1. Ekonomi dengan E1 dianggap normal
2. Ekonomi dengan E5 dianggap tidak normal
3. Ekonomi dengan E7 dianggap tidak normal

KESIMPULAN

Berdasarkan pohon keputusan yang telah dibuat, dapat disimpulkan bahwa atribut perselisihan dan pertengkaran terus menerus, murtad, dan ekonomi adalah faktor-faktor yang berpengaruh terhadap penyebab perceraian di Indonesia. Faktor-faktor ini memiliki nilai *entropy* yang rendah dan nilai *gain* yang lebih besar dibandingkan dengan faktor lainnya. Oleh karena itu, faktor-faktor ini perlu diperhatikan dalam upaya mengurangi tingkat perceraian di Indonesia.

Meskipun pohon keputusan yang dihasilkan relatif sederhana, namun efektif dalam klasifikasi penyebab perceraian antara kategori normal dan tidak normal. Hasil klasifikasi ini berguna untuk berbagai pihak, diantaranya pemerintah, peneliti, dan masyarakat. Untuk pemerintah, klasifikasi ini dapat digunakan untuk mengetahui faktor-faktor yang paling berpengaruh sebagai penyebab perceraian sehingga pemerintah dapat membuat kebijakan yang lebih efektif untuk mengurangi tingkat perceraian di Indonesia. Untuk peneliti, model klasifikasi ini dapat dijadikan sebagai referensi penelitian selanjutnya. Untuk masyarakat, hasil penelitian ini dapat dijadikan pengetahuan dalam upaya mencegah terjadinya penyebab perceraian.

Untuk penelitian selanjutnya, disarankan untuk menggunakan dataset yang lebih besar dan faktor penyebab perceraian yang lebih beragam. Selain itu, dapat juga menggunakan *machine learning* untuk perbandingan hasil klasifikasi. Dengan memperhatikan

hasil penelitian ini, diharapkan angka perceraian dapat berkurang dan kesejahteraan keluarga meningkat.

REFERENSI

- [1] Badan Pusat Statistik, "Statistik Indonesia 2024," Jakarta, 2024. [Online]. Available: <https://www.bps.go.id/publication/2020/04/29/e9011b3155d45d70823c141f/statistik-indonesia-2020.html>
- [2] N. Januari, "MENGGALI AKAR MASALAH: Analisis Kasus Perceraian di Indonesia," *Akad. J. Mhs. Humanis*, vol. 3, no. 3, pp. 120–130, 2023, doi: 10.37481/jmh.v3i3.613.
- [3] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*, 4th Editio. Cambridge: Morgan Kaufmann, 2023.
- [4] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, "Data Mining: Practical Machine Learning Tools and Techniques," *Data Min. Pract. Mach. Learn. Tools Tech.*, pp. 1–621, 2016.
- [5] J. R. Quinlan, "C4.5 Programs for Machine Learning," 1993. [Online]. Available: <http://www.springerlink.com/index/10.1007/BF00993309>
- [6] F. Muzakki, I. Ubaydillah, N. R. Assyiami, and S. Soleha, "Penerapan Algoritma C4.5 Untuk Prediksi Penyakit Jantung Menggunakan Rapidminer," *J. Komput. Antart.*, vol. 2, no. 2, pp. 71–79, 2024, doi: 10.70052/jka.v2i2.304.
- [7] A. I. Silitonga, Z. A. Nabila, C. Rizkia, Z. Lubis, and N. Safitri, "Klasterisasi Gizi Buruk dan Stunting di Provinsi Sumatera Utara Menggunakan K-Means Clustering," vol. 10, no. 2, pp. 13–18, 2024.
- [8] D. Telaumbanua and I. Kurniawati, "Penerapan Algoritma C4.5 Untuk Klasifikasi Kepuasan Pelanggan Pada Jasa Layanan Pengiriman," *JoMMiT J. Multi Media dan IT*, vol. 6, no. 1, pp. 1–6, 2022, doi: 10.46961/jommit.v6i1.524.
- [9] D. Wulandari, N. Lutfiyana, and H. Sumarno, "Metode Algoritma Decision Tree C4.5 Untuk Analisis Kelayakan Kredit Nasabah Pada Bsm Kcp Kemang Pratama," *EVOLUSI J. Sains dan Manaj.*, vol. 7, no. 2, pp. 36–42, 2019, doi: 10.31294/evolusi.v7i2.6757.
- [10] N. Aprilyani, I. Zulfa, and H. Syahputra, "Penerapan Algoritma Decision Tree C4.5 Untuk Model Penentuan Penerima Beasiswa Program Indonesia Pintar (Pip) Studi Kasus Sma Negeri 3 Timang Gajah," *J. Tek. Inform. dan Elektro*, vol. 5, no. 1, pp. 96–109, 2022, doi: 10.55542/jurtie.v5i1.452.
- [11] J. Kumar, *Data Warehouse and Data Mining: Concepts, Techniques, and Real Life Applications*. London: BPB Online, 2024.
- [12] P.-N. Tan, M. Steinbach, A. Karpatne, and V. Kumar, *Introduction to Data Mining*, 2nd Edition. Harlow: Pearson, 2010.
- [13] O. Maimon and L. Rokach, *Data Mining and Knowledge Discovery Handbook*, 2nd Edition. Springer, 2010. doi: 10.1007/978-981-15-8983-6_42.
- [14] A. Fortino, *Data Mining and Predictive Analytics for Business Decisions*. Mercury Learning and Information, 2023.
- [15] B. P. Statistik, "Jumlah Perceraian Menurut Provinsi dan Faktor Tahun 2023," Jakarta, 2023. [Online]. Available: <https://www.bps.go.id/id/statistics-table/3/YVdoU1IwVmITM2h4YzFoV1psWkViRXhqtIZwRFVUMDkjMw==/jumlah-perceraian-menurut-provinsi-dan-faktor.html?year=2023>.

