

## IMPLEMENTASI ALGORITMA C4.5 UNTUK MENENTUKAN TINGKAT PERCERAIAN DI KABUPATEN MALUKU TENGAH

Jama Toyo<sup>1</sup>, Asfin Herianto<sup>2</sup>, Nuru Yarkuran<sup>3</sup>, Syamsul Arif<sup>4</sup>

Program Studi Teknologi Informasi, ITBM Wakatobi, Indonesia<sup>1</sup>  
Program Studi Teknologi Informasi, ITBM Wakatobi, Indonesia<sup>2</sup>  
Program Studi Teknologi Informasi, ITBM Wakatobi, Indonesia<sup>3</sup>  
Program Studi Teknologi Informasi, ITBM Wakatobi, Indonesia<sup>4</sup>

Email: [jamatoyo1703@gmail.com](mailto:jamatoyo1703@gmail.com)<sup>1</sup>, [asfinherianto@gmail.com](mailto:asfinherianto@gmail.com)<sup>2</sup>,  
[nuruyarkuran.01@gmail.com](mailto:nuruyarkuran.01@gmail.com)<sup>3</sup>, [sa200620@gmail.com](mailto:sa200620@gmail.com)<sup>4</sup>

### ABSTRAK

Perceraian merupakan salah satu permasalahan dalam kehidupan berkeluarga sehingga masyarakat pada umumnya beranggapan bahwa penyebab utama perceraian adalah faktor ekonomi dan sosial, misalnya kesenjangan ekonomi, kekerasan dalam rumah tangga (KDRT), hingga perselingkuhan. Namun tidak menutup kemungkinan ada faktor-faktor dominan lainnya yang mungkin belum teridentifikasi. Sehingga penelitian ini membantu untuk memberikan informasi kepada pihak terkait dalam membijaki persoalan perceraian di Kabupaten Maluku Tengah dengan menggunakan rules pada C4.5. Dalam penelitian ini kami menggunakan Algoritma C4.5 dalam 11 variabel (umur suami, umur istri, Pendidikan suami, pendidikan Istri, pekerjaan suami, pekerjaan istri, pendapatan suami, Pendapatan Istri, Lama pernikahan, jumlah anak, Indikator Perceraian. Dari hasil analisis ditemukan 40 (empat puluh) pola untuk mengidentifikasi potensi akan terjadinya perceraian. Dari perhitungan Algoritma C.45 maka variabel indikator Penyebab adalah indikator utama untuk menentukan status cerai atau tidak cerai dengan nilai gain tertinggi 0.0069, kemudian nilai gain terendah terdapat pada variabel pendapatan dan pekerjaan suami gain 0.0498 dan -0.0124..

**Kata Kunci:** Implementasi, Decision Tree C4.5, Data Mining.

### ABSTRACT

*Divorce is such a problem with family life that people in general feel that the leading cause of divorce is economic and social factors such as the inequality of domestic violence and therefore of infidelity. But it does not rule out other possible dominant factors that may yet be identified. So this study is helpful in providing information to those associated with divorce in central Maluku district by taking rules at C4.5. In this study we're using a C4.5 algorithm in 11 variables (husband's age, wife's age, a husband's education, wife's education, husband's occupation, wife's work, this income, wife's income, old marriage, number of children, divorce indicators). analysis found 40 patterns to identify the potential for divorce. Based on the C4.5 algorithm calculations, it is the main indicator for determining the rate of divorce or not for divorce with the highest gain value 0.0069, and low game value found on the income and job variables gain 0.0498 and -0.0124*

**Keywords:** Implementasi, Decision Tree C4.5, Data Mining

## 1. PENDAHULUAN

Berdasarkan data Direktorat Jenderal Badan Peradilan Agama (Badilag) yang dirilis pada laman *Hukumonline.com*, perkara putusan (*inkracht*) perceraian di 29 Pengadilan Tinggi Agama seluruh Indonesia dalam kurun waktu tiga tahun terakhir (2015-2017) menunjukkan peningkatan yang signifikan. Kasus perceraian pada tahun 2016 berjumlah 0,38% sedangkan pada tahun 2017 perkara perceraian meningkat 0,98 %. Tercatat pada tahun. Meskipun kasus perceraian di kota Ambon cenderung menurun 1,82 % pada tahun 2017, angka ini masih sangat tinggi.

Meningkatnya kasus perceraian di kota Ambon dengan penyebab utama perceraian adalah faktor ekonomi dan sosial, misalnya kesenjangan ekonomi, kekerasan dalam rumah tangga, hingga perselingkuhan. Namun tidak menutup kemungkinan ada faktor-faktor dominan lainnya yang mungkin belum teridentifikasi.

Penelitian tentang kasus perceraian telah dilakukan beberapa peneliti yaitu *Perbandingan Metode Naïve Bayes dan K-Nearest Neighbort untuk Prediksi Perceraian* Untuk mengetahui seberapa banyak perceraian disetiap bulannya. Tingginya Angka Cerai Gugat (Khulu') di Indonesia; Analisis Kritis Terhadap Penyebab dan Alternatif Solusi Mengatasinya. Sehingga perlu adanya pengkajian terlebih dahulu mengenai variabel-variabel apa saja yang berpotensi akan terjadinya perceraian.

Tujuan penelitian ini adalah untuk klasifikasikan variabel-variabel apa saja yang berpotensi akan terjadinya perceraian di Pengadilan Agama Maluku Tengah. Peneliti ini akan mengelompokkan variabel-variabel apa saja yang berpotensi akan terjadinya perceraian menggunakan

Algoritma C4.5 untuk menguji hasil akurasi berpotensi perceraian menggunakan *precision, recall dan acurasi*.

## 2. LANDASAN TEORI

Berdasarkan penelitian Irma Apriliana Dahlia, Muhammad Irfan dan Wisnu, 2018 *Perbandingan Metode Naïve Bayes dan K-Nearest Neighbort untuk Prediksi Perceraian (Studi kasus : Pengadilan Agama Cimahi)*, Untuk mengetahui seberapa banyak perceraian disetiap bulannya [1] Isnawati Rais, 2014 Tingginya Angka Cerai Gugat (Khulu') di Indonesia; Analisis Kritis Terhadap Penyebab dan Alternatif Solusi Mengatasinya. Meningkatkan jumlah kasus perceraian dari tahun ke tahun, khususnya cerai gugat, baik di tingkat nasional maupun lokal di Jakarta Selatan. [2] Dari penelitian terdahulu maka peneliti merasa perlu adanya pengkajian terlebih dahulu mengenai variabel-variabel apa saja yang berpotensi akan terjadinya perceraian

## 3. METODE PENELITIAN

### 3.1. Metode pengumpulan data

Metode pengumpulan data yang dilakukan peneliti yaitu observasi langsung yang telah dilakukan di lokasi penelitian yaitu pada Pengadilan Agama Maluku Tengah.

*Data mining* adalah sebuah istilah yang digunakan untuk menggambarkan penemuan ilmu pengetahuan dalam bidang database, sebuah bidang analisis informasi yang mencari pola tersembunyi dalam sekelompok data yang dapat digunakan untuk memprediksi perilaku masa depan [3]. Data Kasus Perceraian pada tahun 2014 sampai dengan 2018, yang terdiri dari tahun 2014 jumlah kasus sebanyak 111 perkara, 2015, sebanyak 162 perkara,

2016, sebanyak 124 perkara dan tahun 2017 sebanyak 106 perkara, perkara tahun 2018 sebanyak 101 perkara, maka total perkara keseluruhan adalah sebanyak 604 perkara. Terdiri dari data training 564 data dan data testing 40 data.

### 3.2. Algoritma C45

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti Structured Query Language untuk mencari record pada kategori tertentu.[4] Algoritma C4.5 mempunyai input berupa *training samples* dan *samples*. *Training samples* berupa data contoh yang akan digunakan untuk membangun sebuah *tree* yang telah diuji kebenarannya. Sedangkan *samples* merupakan *field-field* data yang nantinya akan kita gunakan sebagai parameter dalam melakukan klasifikasi data [5].

Sebuah obyek yang diklasifikasikan dalam pohon harus dites nilai *Entropy* -nya. [5]  $(S) = \sum_{i=1}^n -p_i \log_2 p_i = 1$  (1). Dari nilai *Entropy* tersebut kemudian dihitung nilai *information gain* (IG) masing-masing atribut dengan menggunakan Persamaan  $Gain(S_i) = entropy(S) - \sum_{i=1}^n |S_i|/|S| * entropy(S_i)$   $ni = 1(2)$ .

## 4. HASIL DAN PEMBAHASAN

### 4.1. Proses Perhitungan

Entropy Total

$$= \frac{-13}{564} * \log_2 \frac{13}{564} + \left( \frac{-551}{564} * \log_2 \frac{551}{564} \right) = 0.03080137 + (-0.284197291)$$

$$= -0.253395917$$

Keterangan :

S : Himpunan kasus (S). = 77

n : Jumlah kasus (S<sub>1</sub>) No = 4

Pi : Proporsi (S<sub>2</sub>) = 73

Usia Istri (Remaja) (S) Total = ?

Usia Istri (Remaja)

$$= \frac{-4}{77} \log_2 \frac{4}{77} + \left( \frac{-73}{77} * \log_2 \frac{73}{77} \right) = 0.05108575 + (-0.263427734) = -0.212341982$$

$$\text{Dewasa} = \frac{-9}{424} \log_2 \frac{9}{424} + \left( \frac{-415}{424} \log_2 \frac{415}{424} \right) = 0.02912462 + (-0.285520231) = -0.256395608$$

$$\text{Lansia} = \frac{-0}{63} \log_2 \frac{0}{63} + \left( \frac{-63}{63} \log_2 \frac{63}{63} \right) = 0 + (-0.301029996) = -0.301029996$$

$$Gain(S) = Entropy(s) - \sum_{i=1}^n \frac{S_i}{S} * Entropy(S_i)$$

Gain

$$= -0.253395917 - \left( \frac{77}{564} * (0.212341982) + \frac{424}{564} * (-0.256395608) + \frac{63}{564} * (-0.301029996) \right) = (-0.02898995 + (-0.19275131) + (-0.03362569)) = -0.254686466 - (-0.25536695) = 0.001971.$$

Tabel 1 Usia Perceraian Istri dan Suami

No	Rentang Umur	Keterangan
11	17-25	Remaja
22	26-45	Dewasa
33	46-65	Lansia
44	>65	Manula

Table 2. terdiri dari 3 kategori yaitu Pendidikan dasar Pendidikan Menengah dan Pendidikan Tinggi berdasarkan pada

Undang-undang No.20 tahun 2003 yang terdapat dalam kajian teori [6]

Tabel 2. Pendidikan dan pengkategorian

No	Sekolah	Kategori
1	SD, SMP	Pendidikan Dasar
2	SMA	Pendidikan Menengah
3	D2-S3	Pendidikan Tinggi

Variabel pekerjaan terdiri dari tiga kategori [7] yaitu. Status Tinggi, Menengah dan Rendah, sedangkan golongan pekerjaan, peneliti hanya mengambil 9 kelompok, berdasarkan pada ketentuan KBJI (2002), yang terdapat dalam pembahasan tentang pekerjaan pada kajian teori, seperti terdapat pada tabel 3 berikut ini :

Tabel 3 Kelompok Pekerjaan.

No	Status Pekerjaan
1.	Status Tinggi
2.	Status Menengah
3.	Status Rendah

Pengkategorian rentang Pendapatan pada tabel 3.4 berdasarkan data Statistik (BPS) 2011 [7] pada kajian teori dapat digolongkan menjadi 4 yaitu

Tabel 4. Pendapatan dan Pengkategorian [8]

No	Rentang	Kategori
1	> 3.500.000	Pendapatan Sangat Tinggi
2	2.500.000 – 3.500.000	Pendapatan Tinggi
3	1.500.000 – 2.500.000	Pendapatan Sedang
4	< 1.500.000	Pendapatan Rendah

Berdasarkan pada kajian teori, variabel lama pernikahan terdiri dari 4 kategori [1]. yaitu : Berisi tentang analisis

penelitian dan temuan-temuan terbaru yang ditemukan dalam penelitian.

Tabel 5. Variabel Lama Pernikahan

No	Lama Pernikahan	Kategori
1	1-5	Sangat Cepat
2	6-11	Cepat
3	12-17	Lama
4	> 18	Sangat Lama

Pengkategorian tabel 6 berdasarkan pada kajian teori, maka jumlah anak ternagi menjadi tiga yaitu [9] :

Tabel 6. Pengelompokkan Jumlah Anak

No	Jumlah Anak	Kategori
1	1	Sedikit
2	2	Sedang
3	>3	Banyak

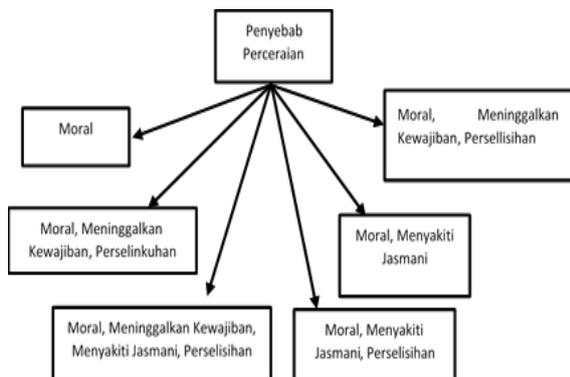
Pengclasteran indikator penyebab perceraian terdiri dari 8 berdasarkan pada KHI yang terdapat pada kajian teori. [10]

Tabel.7 Indikator Penyebab Perceraian

No	Indikator Penyebab Perceraian
1	Moral
2	Meninggalkan Kewajiban
3	Hukum Pidana
4	Menyakiti Jasmanai
5	Cacat Biologis
6	Perselisihan
7	Taklik Talak
8	Murtad

Dari tabel 1-7 terlihat bahwa ada sebelas variabel berpotensi akan terjadinya perceraian karena faktor ketidakcocokan dan tiap-tiap variabel memiliki nilai gain

yang dihasilkan algoritma C45 di antaranya Usia Istri 0.0020, Pendidikan Istri 0.0003, pekerjaan Istri 0.0010, Pendapatan Istri 0.0028, Usia Suami 0.0022, Pendidikan suami 0.0019, Pendapatan Suami 0.0028, Lama Pernikahan 0.0002, Jumlah anak 0.0004, Penyebab 0.0069. Maka jumlah gain dari keseluruhan variabel dengan menggunakan Algoritma C45, nilai gain yang tertinggi dijadikan root yaitu variabel Penyebab, sehingga variabel tersebut menjadi root node pada kasus perceraian di pengadilan Agama Masohi, seperti di gambarkan pada gambar 1



Gambar. 1 Root Node

Gamabra 1 menggambarkan bahwa yang menjadi root node pada penelitian ini adalah indikator penyebab perceraian,

### 3.3. Pengujian Data Testing

Pengujian data kasus berdasarkan rules dengan menggunakan algoritma C45, data testing yang dugnakan sebanyak 40 data yang disediakan, tidak termasuk dalam data training.

### 1. Algoritma C45

Proses pengujian yang dilakukan dengan menggunakan data testing pada sebanyak 40 data dengan menggunakan algoritma C45, maka hasil yang di peroleh berdasarkan rules dengan menghitung presentase nilai cerai dan tidak cerai, sebagai berikut :

Table 8. hasil pengujian data testing

NO	CLASS PREDIKSI	CERAI	TIDAK CERAI	KESESUAIAN
1	Cerai	2.000%	0.100%	Cerai
2	Cerai	0.000%	0.000%	Tidak Cerai
3	Cerai	0.000%	0.000%	Tidak Cerai
4	Cerai	3.000%	0.200%	Cerai
5	Cerai	2.326%	0.000%	Cerai
6	Cerai	2.726%	0.000%	Cerai
7	Cerai	2.626%	0.000%	Cerai
8	Tidak Cerai	0.000%	0.000%	Tidak Cerai
9	Tidak Cerai	0.000%	0.000%	Tidak Cerai
10	Cerai	2.926%	0.000%	Cerai
11	Cerai	2.426%	0.000%	Cerai
12	Cerai	2.746%	0.200%	Cerai
13	Cerai	2.126%	0.000%	Cerai
14	Cerai	3.326%	0.000%	Cerai
15	Cerai	3.446%	1.200%	Cerai
16	Cerai	2.646%	0.100%	Cerai

17	Cerai	2.626%	0.000%	Cerai
18	Cerai	2.826%	0.000%	Cerai
19	Cerai	2.225%	0.000%	Cerai
20	Cerai	2.926%	0.000%	Cerai
21	Cerai	0.000%	0.000%	Cerai
22	Cerai	2.526%	0.4000%	Cerai
23	Cerai	2.826%	0.1000%	Cerai
24	Cerai	2.926%	0.000%	Cerai
25	Cerai	2.246%	0.000%	Cerai
26	Cerai	2.826%	0.1000%	Cerai
27	Cerai	2.426%	0.5000%	Cerai
28	Cerai	2.426%	0.5000%	Cerai
29	Cerai	2.626%	0.3000%	Cerai
30	Cerai	3.826%	0.100%	Cerai
31	Cerai	0.000%	0.000%	Cerai
32	Cerai	0.000%	0.000%	Cerai
33	Cerai	0.000%	0.000%	Cerai
34	Cerai	0.000%	0.000%	Cerai
35	Cerai	0.000%	0.000%	Cerai
36	Cerai	0.000%	0.000%	Cerai
37	Cerai	0.000%	0.000%	Cerai
38	Cerai	0.000%	0.000%	Cerai
39	Cerai	2.726	0.2	Cerai
40	Cerai	2.626	0.3	Cerai

Berdasarkan pada data testing maka yaitu data pada kasus 8 dan 9 terdapat perbedaan dimana, pada data aslinya tidak bercerai namun pada data

testingnya bercerai. Oleh karena itu dibuatlah tabel perbandingan dan tabel persamaan untuk menghitung precision, recall dan accuracy dari proses perhitungan algoritma C45

Tabel. 9 Precision, Recall dan Akurasi

Nilai Prediksi	Kategori	
	Cerai	36 (True Positive)
Tidak Cerai	0 False Positif)	2 (True Negatif)

Proses pengujian yang dilakukan dengan menggunakan data testing pada tabel 8 dan berdasarkan hasil perhitungan entropy dan gain pada data training serta rules yang di hasil dari pohon keputusan yang telah dibuat. Proses pengujian juga dilakukan dengan menghitung nilai precision, recall dan accuracy.

Rumus :

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Keterangan :

TP : True Positif

TN : True Negatif

FP : False Positif

FN : False Negatif

$$Precision = \frac{36}{36 + 0} \times 100 = 100 \%$$

$$Recall = \frac{36}{36 + 2} \times 100 = 94 \%$$

$$\text{Accuracy} = \frac{36 + 2}{36 + 2 + 0 + 2} \times 100 \\ = 95 \%$$

Hasil perhitungan algoritma C45 menunjukkan bahwa nilai precision sebesar 100%, Nilai Recall 94,73% dan Nilai Accuracy 95 %.

## 5. SIMPULAN DAN SARAN

### 5.1. Simpulan

Dari Dari penelitian yang telah dilakukan dapat di simpulkan beberapa hal yaitu :

1. Algoritma C45 memberikan informasi dalam bentuk nilai presentase akurasi algoritma dalam proses pengolahan data kasus, dimana hasil implementasi yang di peroleh dengan menggunakan algoritma C45 berjumlah 95%.
2. Perhitungan Algoritma C45 dalam 11 variabel mempunyai potensi akan terjadinya perceraian karena faktor ketidak cocokan, namun berdasarkan data kasus pada Pengadilan Agama Masohi, dengan algoritma C45 maka nilai gain tertinggi terdapat pada variabel indikator Penyebab dengan jumlah gainnya 0.0069, variabel pendapatan dan pekerjaan suami dengan jumlah gain 0.0498 dan -0.0124. variabel indikator Penyebab berpotensi rentang terjadinya perceraian, ketika kita di hadapkan dengan indikator penyebab maka solusinya adalah menyelesaikan dengan cara yang baik dan tepat. Sedangkan variabel **pendapatan dan pekerjaan suami** berpotensi akan terjadinya perceraian jika tidak dimilikinyamaka solusinya seorang suami harus memiliki pekerjaan karena pendapatan didapatkan dari pekerjaan suami, meskipun variabel

yang lain berpotensi terjadinya perceraian namun solusi ini berdasarkan data pada Pengadilan Agama Masohi.

### 5.2. Saran

Pada penelitian ini, peneliti hanya menggunakan satu Algoritma C.45 diharapkan pada peneliti selanjutnya menggunakan Algoritma yang KNN sehingga bisa mengetahui tentang akurasi tiap algoritma.

## DAFTAR PUSTAKA

- [1]. Irma Apriliana Dahlia, Mohammad Irfan dan Wisnu Uriawan, 2018, *Perbandingan Metode Naïve Bayes dan K-Nearest Neighbour untuk Prediksi Perceraian (Studi kasus : Pengadilan Agama Cimahi)*, Jurnal Teknik Informatika, Fakultas Sains dan UIN SGD Bandung, INSIGHT Volume 1 No. 1 : 182-187
- [2]. Isnawati Rais, 2014, *Tingginya Angka Cerai Gugat (Khulu') di Indonesia; Analisis Kritis Terhadap Penyebab dan Alternatif Solusi Mengatasinya*, Jurnal Fakultas Syariah dan Hukum UIN Syarif Hidayatullah Jakarta, Al-Adalah Vol.XII, No. 1 Juni 2014.
- [3]. Ni Komang Sri Julyyantari, I Ketut dedy Suryawan, 2014, *Data Mining Prestasi, Akademik Dengan Naïve Bayes Berdasarkan Atribut Importance (AI)*, Jurnal Sistem dan Informatika STMIK STIKOM Bali, Vol. 9 No.1, Jl. Raya Puputan No. 68 Renon, Denpasar – Bali
- [4]. Eka Pandu Cynthia, Edi Ismanto, 2018, *Metode Decision Tree Algoritma C4.5 dalam Mengklasifikasi data Penjualan Bisnis Gerai makanan Cepat Saji*, Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK) Volume (3) Juli 2018, pp 1-13 ISSN: 2527-5771/EISSN: 2549-7839.

- [5]. Esty Purwaningsih, 2016, *Seleksi Mobil Berdasarkan Fitur dengan Komparasi Metode Klasifikasi Neural Network, Support Vektor Machine, dan Algoritma C4.5*, Program Studi Manajemen Informatika AMIK BSI Tangerang, Jurnal Pilar Nusa Mandiri Vol.XII, No.2 September 2016.
- [6]. Undang-undang RI No. 20 Tahun 2003, file Pdf download. [https://kelembagaan.ristekdikti.go.id/wp-content/uploads/2016/08/UU\\_no\\_20\\_th\\_2003.pdf](https://kelembagaan.ristekdikti.go.id/wp-content/uploads/2016/08/UU_no_20_th_2003.pdf). 07:43 tgl 23 juni 2019.
- [7]. KBJI 2002, katalog BPS : 3420 download file pdf.
- [8]. Lia Kurniawati, dkk, Hubungan antara tingkat pendidikan, Status pekerjaan dan tingkat pendapatan dengan usia perkawinan pertama anita Di kelurahan kotalama kecamatan kedungkandang kota malang, Jurnal Preventia, Vol 2 No 1 Juni 2017, Fakultas Ilmu Keolahragaan Universitas Negeri Malang.
- [9]. Hartoyo, melly latifah, sri rahayu mulyani, 2011, studi nilai anak, jumlah anak yang diinginkan, dan keikutsertaan orang tua dalam program kb Departemen Ilmu Keluarga dan Konsumen, Fakultas Ekologi Manusia, Institut Pertanian Bogor,
- [10]. Ahmad Bastomi, 2010, *Tinjauan Kompilasi Hukum Islam dan PP No.9 Tahun 1975 terhadap pelaksanaan Pencatatan Perceraian di Kantor Urusan Agama Kec. Gurah Kab. Kediri*, Skripsi Jurusan Ahwal al-Syaksiyyah Fakultas Syari'ah Institut Agama Islam Negeri Sunan Ampel Surabaya.