

KLASIFIKASI STATUS PERKARA PERCERAIAN MENGUNAKAN ALGORITMA C4.5 PADA PENGADILAN AGAMA CIKARANG

Naila Hanin Ayudie¹⁾, Yulan Febriyanti²⁾, Elkin Rilvani³⁾, Erna Apriani⁴⁾

^{1,2,4} Program Studi Bisnis Digital, Universitas Pelita Bangsa

³ Program Studi Teknik Informatika, Universitas Pelita Bangsa

^{1,2,3,4} Jl. Inspeksi Kalimalang No.9, Cibatu, Cikarang Sel., Kabupaten Bekasi, Jawa Barat

Email: ¹nailahaninayudie14@gmail.com, ²Yulanyuliani82@gmail.com,

³elkin.rilvani@pelitabangsa.ac.id, ⁴ernaapriani@pelitabangsa.ac.id

Abstract

The increasing number of divorce cases has resulted in a large volume of data in the Case Tracking Information System (SIPP) of the Cikarang Religious Court, making manual data analysis less effective. This study aims to classify the status of divorce cases using the C4.5 algorithm based on data from the SIPP of the Cikarang Religious Court in 2025. The research employed a quantitative approach using the CRISP-DM methodology, which consists of *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, and *Evaluation*. The dataset consisted of 4,554 divorce case records, with Case Classification and Processing Time as predictor variables and Case Status as the target variable. The evaluation results using *Cross Validation* showed that the proposed model achieved an accuracy of $88.63\% \pm 0.83\%$. The resulting decision tree indicated that Processing Time was the most influential attribute in determining the classification of case status. However, the model's performance on minority classes was relatively low due to class imbalance in the dataset. Overall, the C4.5 algorithm proved effective in classifying divorce case status and can support data analysis and decision-making in managing divorce case data at the Cikarang Religious Court.

Keywords: C4.5, Data Mining, Decision Tree, Case Status, SIPP.

Abstrak

Peningkatan jumlah perkara perceraian menyebabkan data pada Sistem Informasi Penelusuran Perkara (SIPP) Pengadilan Agama Cikarang semakin besar sehingga analisis secara manual menjadi kurang efektif. Penelitian ini bertujuan mengklasifikasikan status perkara perceraian menggunakan algoritma C4.5 berdasarkan data SIPP Pengadilan Agama Cikarang Tahun 2025. Penelitian menggunakan pendekatan kuantitatif dengan metodologi CRISP-DM yang meliputi *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, dan *Evaluation*. Dataset yang digunakan terdiri dari 4.554 data perkara perceraian dengan atribut Klasifikasi Perkara dan Lama Proses sebagai variabel prediktor serta Status Perkara sebagai variabel target. Hasil evaluasi menggunakan *Cross Validation* menunjukkan bahwa model memperoleh tingkat akurasi sebesar $88,63\% \pm 0,83\%$. Pohon keputusan yang dihasilkan menunjukkan bahwa atribut Lama Proses merupakan atribut yang paling berpengaruh dalam proses klasifikasi status perkara. Meskipun demikian, kemampuan model dalam mengenali kelas minoritas masih terbatas akibat ketidakseimbangan distribusi data. Secara keseluruhan, algoritma C4.5 terbukti mampu menghasilkan model klasifikasi yang efektif dan dapat dimanfaatkan sebagai pendukung analisis serta pengambilan keputusan dalam pengelolaan data perkara perceraian di Pengadilan Agama Cikarang.

Kata Kunci: C4.5, Data Mining, Decision Tree, Status Perkara, SIPP

1. PENDAHULUAN

Perceraian merupakan upaya untuk melepaskan ikatan suami dan istri dari suatu perkawinan yang disebabkan oleh alasan tertentu. Perceraian terjadi karena sudah tidak adanya jalan keluar (*dissolution marriage*) [1]. Jumlah kasus perceraian yang tercatat tiap tahun terus meningkat sehingga menghasilkan volume data yang besar. Data ini disimpan dalam Sistem Informasi Penelusuran Perkara (SIPP) dan dapat dimanfaatkan untuk memperoleh informasi penting yang mendukung proses pengambilan keputusan [2]. Data perceraian yang tersimpan di SIPP tidak hanya memuat informasi mengenai jenis perkara, tetapi juga status perkara dan lama proses penyelesaiannya. Besarnya jumlah data menyebabkan proses analisis secara manual menjadi kurang efektif sehingga diperlukan teknik pengolahan data yang mampu menemukan pola dari data tersebut [3].

Salah satu pendekatan untuk mengolah data perkara adalah data mining, terutama teknik klasifikasi. Teknik klasifikasi berfungsi mengelompokkan data ke dalam kelas-kelas tertentu berdasarkan pola yang dipelajari dari data historis. Algoritma C4.5 merupakan contoh metode klasifikasi yang dapat membentuk model pohon keputusan yang mudah dipahami dan diinterpretasikan [4].

Walaupun banyak penelitian telah menggunakan algoritma C4.5 untuk berbagai kasus klasifikasi [5], penerapan algoritma tersebut untuk mengklasifikasikan status perkara perceraian berdasarkan data SIPP Pengadilan Agama Cikarang Tahun 2025 masih belum banyak dilakukan. Oleh karena itu, penelitian ini dilakukan untuk membangun model klasifikasi status perkara menggunakan algoritma C4.5 sehingga dapat memberikan informasi mengenai pola penyelesaian perkara berdasarkan atribut yang digunakan.

2. METODE PENELITIAN

2.1 Metode Penelitian

Penelitian ini menggunakan metode penelitian kuantitatif dengan pendekatan data mining. Pendekatan data mining dipilih karena mampu mengolah data dalam jumlah besar untuk menemukan pola serta membangun model klasifikasi yang dapat mendukung proses pengambilan keputusan. Pada penelitian ini, algoritma yang digunakan adalah C4.5, yaitu salah satu algoritma klasifikasi yang menghasilkan model dalam bentuk *decision tree* (pohon keputusan) sehingga mudah dipahami dan diinterpretasikan [6].

Metodologi yang digunakan dalam penelitian ini adalah *Cross Industry Standard Process for Data Mining* (CRISP-DM). Metodologi ini dipilih karena memiliki tahapan yang terstruktur dalam proses pengembangan model data mining, dimulai dari memahami masalah penelitian hingga mengevaluasi model yang dihasilkan. Tahapan CRISP-DM yang dilaksanakan dalam penelitian ini meliputi *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, dan *Evaluation*. Tahap *Deployment* tidak dilakukan karena penelitian ini hanya berfokus pada pembangunan dan evaluasi model klasifikasi [7].

2.2 Objek Penelitian

Objek penelitian ini berupa data perkara perceraian yang diperoleh dari Sistem Informasi Penelusuran Perkara (SIPP) Pengadilan Agama Cikarang Tahun 2025. Data tersebut memuat informasi mengenai proses penyelesaian perkara perceraian yang digunakan sebagai dasar dalam pembangunan model klasifikasi. Fokus penelitian ini tertuju pada mengklasifikasikan status perkara perceraian menggunakan algoritma C4.5 sehingga diperoleh pola hubungan antara karakteristik perkara dengan status perkara yang dihasilkan.

2.3 Dataset Penelitian

Dataset penelitian terdiri atas 4.554 data perkara perceraian tahun 2025 yang telah melalui proses seleksi dan pembersihan data sehingga siap digunakan pada tahap pemodelan.

Tabel 1. Atribut *Dataset*

Atribut	Peran
Klasifikasi Perkara	Prediktor
Lama Proses	Prediktor
Status Perkara	Label

2.4 Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma klasifikasi yang menghasilkan model berupa pohon

keputusan (*decision tree*). Algoritma ini membentuk pohon keputusan dengan memilih atribut yang memiliki nilai *Information Gain* tertinggi sebagai akar pohon (*root node*) [8]. Proses tersebut dilakukan secara berulang hingga seluruh data berhasil dikelompokkan ke dalam kelas tertentu.

Perhitungan *Information Gain* diawali dengan menghitung nilai *Entropy* untuk mengetahui tingkat ketidakpastian suatu kumpulan data. Semakin kecil nilai *Entropy*, maka semakin homogen data tersebut. Berikut rumus *Entropy*.

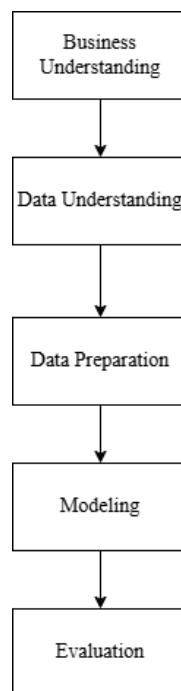
$$Entropy(S) = - \sum_{i=1}^n p_i \log_2(p_i) \quad (1).$$

Selanjutnya, nilai *Information Gain* dihitung untuk menentukan atribut terbaik yang akan digunakan sebagai node pada pohon keputusan. Berikut rumus *Information Gain*.

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

2.5 Tahapan Penelitian

Penelitian ini menerapkan metodologi CRISP-DM (*Cross Industry Standard Process for Data Mining*) yang meliputi lima tahapan utama yang digunakan dalam studi ini, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, dan *Evaluation*. Tahap *Deployment* tidak dilaksanakan karena fokus penelitian ini hanya berfokus pada pembuatan dan penilaian model klasifikasi.



Gambar 1. Tahapan Menggunakan Metodologi CRISP-DM

a. *Business Understanding*

Tahap ini bertujuan mengidentifikasi permasalahan penelitian, yaitu belum tersedianya model yang mampu mengklasifikasikan status perkara perceraian berdasarkan karakteristik data pada SIPP Pengadilan Agama Cikarang. Hasil dari tahap ini adalah penetapan tujuan penelitian, yaitu membangun model klasifikasi menggunakan algoritma C4.5.

b. *Data Understanding*

Dataset yang digunakan berasal dari Sistem Informasi Penelusuran Perkara (SIPP) Pengadilan Agama Cikarang Tahun 2025 yang berjumlah 4.554 data perkara. Sebelum digunakan dalam proses pemodelan, dilakukan pemeriksaan distribusi data pada setiap atribut untuk mengetahui karakteristik dataset. Analisis ini bertujuan memastikan bahwa atribut yang digunakan memiliki kualitas data yang baik serta mampu merepresentasikan kondisi sebenarnya. Selain itu, dilakukan identifikasi terhadap kemungkinan terjadinya ketidakseimbangan distribusi kelas (*class imbalance*) yang dapat memengaruhi hasil klasifikasi.

c. *Data Preparation*

Tahap data preparation tidak hanya berfokus pada pembersihan data, tetapi juga memastikan konsistensi nilai atribut yang digunakan dalam proses klasifikasi. Data yang memiliki nilai kosong atau tidak lengkap dihapus agar tidak memengaruhi proses pembentukan *decision tree*. Selanjutnya atribut yang tidak memiliki kontribusi terhadap proses klasifikasi tidak digunakan sehingga model hanya memanfaatkan atribut yang relevan. *Dataset* kemudian dikonversi ke format *CSV* agar kompatibel dengan perangkat lunak Altair AI Studio.

d. *Modeling*

Pada tahap ini meliputi pembangunan model klasifikasi menggunakan algoritma C4.5. *Dataset* diimpor ke Altair AI Studio, kemudian atribut Status Perkara diterapkan sebagai label sementara atribut Klasifikasi Perkara serta Lama Proses berperan sebagai prediktor. Selanjutnya dilakukan pelatihan model menggunakan data latih dan pengujian dengan teknik validasi untuk menghasilkan model klasifikasi yang optimal.

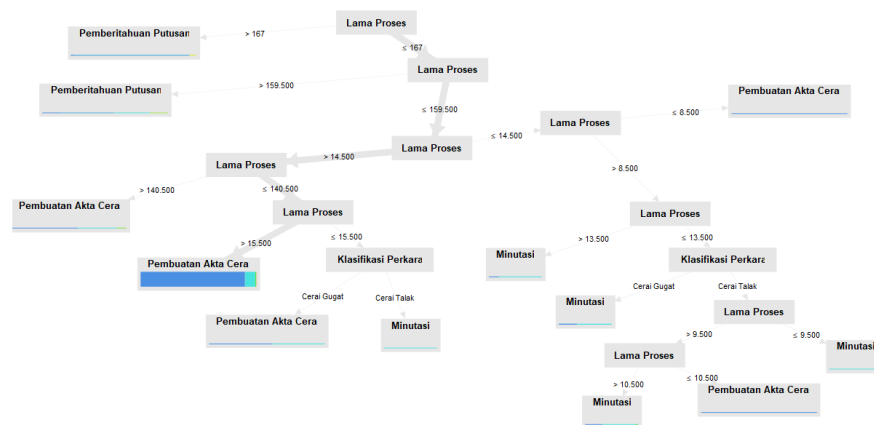
e. *Evaluation*

Tahap evaluasi bertujuan menilai performa model klasifikasi yang telah dibuat. Penilaian dilakukan dengan *Confusion Matrix*, *Accuracy*, *Precision*, dan *Recall*. Nilai-nilai tersebut digunakan untuk mengevaluasi sejauh mana algoritma C4.5 mampu mengklasifikasikan status perkara perceraian pada data SIPP Pengadilan Agama Cikarang.

3. HASIL DAN PEMBAHASAN

3.1 Pembentukan Model Algoritma C4.5

Proses pemodelan dilakukan menggunakan perangkat lunak Altair AI Studio dengan algoritma C4.5. *Dataset* diimpor ke aplikasi, kemudian data yang sudah melalui tahap *data preparation* digunakan untuk membangun model klasifikasi dengan atribut Klasifikasi Perkara dan Lama Proses sebagai atribut prediktor, sedangkan Status Perkara sebagai atribut label. Hasil proses pemodelan menghasilkan sebuah pohon keputusan (*decision tree*) yang menggambarkan aturan-aturan klasifikasi status perkara [9].



Gambar 2. Pohon Keputusan

Berdasarkan **Gambar 2**, diketahui bahwa atribut Lama Proses menjadi *root node* atau akar pohon keputusan. Hal ini menunjukkan bahwa atribut tersebut memiliki nilai *information gain* tertinggi dibandingkan atribut lain, sehingga dipilih sebagai atribut pertama dalam proses pembentukan pohon keputusan. Dengan demikian, lama proses penyelesaian perkara menjadi faktor yang paling berpengaruh dalam menentukan klasifikasi status perkara.

Selanjutnya, atribut Lama Proses kembali muncul pada beberapa cabang pohon keputusan dengan batas nilai tertentu, seperti ≤ 167 hari, $\leq 159,5$ hari, $\leq 140,5$ hari, $\leq 15,5$ hari, $\leq 14,5$ hari, $\leq 13,5$ hari, $\leq 10,5$ hari, dan $\leq 9,5$ hari. Penggunaan atribut yang sama pada beberapa tingkat pohon merupakan karakteristik algoritma C4.5, yaitu algoritma akan menghitung kembali nilai *information gain* pada setiap subset data hasil pembagian sebelumnya. Apabila atribut yang sama masih memiliki nilai gain tertinggi, maka atribut tersebut akan digunakan kembali sebagai node berikutnya.

Selain atribut Lama Proses, model juga menggunakan atribut Klasifikasi Perkara sebagai atribut lanjutan pada beberapa cabang pohon keputusan. Atribut ini digunakan untuk membedakan jenis perkara, yaitu Cerai Gugat dan Cerai Talak, sehingga model dapat menghasilkan prediksi status perkara yang lebih spesifik.

Berdasarkan struktur pohon keputusan yang terbentuk, model menghasilkan beberapa kelas prediksi, yaitu Pembuatan Akta Cerai, Pemberitahuan Putusan, dan Minutasi. Setiap daun (*leaf node*) pada

pohon keputusan menunjukkan hasil akhir klasifikasi berdasarkan kombinasi nilai atribut yang telah diproses oleh algoritma C4.5.

Pada beberapa node masih terlihat adanya lebih dari satu warna yang menunjukkan bahwa data pada node tersebut belum sepenuhnya homogen. Kondisi ini menunjukkan bahwa masih terdapat lebih dari satu kelas status perkara dalam node tersebut, sehingga algoritma menetapkan kelas dengan jumlah data terbanyak sebagai hasil prediksi. Hal tersebut merupakan kondisi yang umum terjadi pada proses pembentukan *decision tree*, terutama ketika distribusi data antar kelas tidak seimbang.

Hasil pembentukan *decision tree* menunjukkan bahwa atribut Lama Proses memiliki pengaruh yang lebih besar dibandingkan atribut Klasifikasi Perkara dalam proses mengklasifikasi status perkara perceraian. Semakin lama atau semakin singkat waktu penyelesaian suatu perkara, maka semakin besar pengaruhnya terhadap penentuan status perkara yang diprediksi oleh model. Oleh karena itu, atribut Lama Proses menjadi atribut utama dalam proses pengambilan keputusan pada model klasifikasi yang dihasilkan.

3.2 Evaluasi Model

Evaluasi model dilakukan menggunakan *Confusion Matrix*, *Accuracy*, dan *Precision* untuk mengetahui kinerja algoritma C4.5 dalam mengklasifikasikan status perkara perceraian pada data SIPP Pengadilan Agama Cikarang Tahun 2025 [10].

Dari hasil pengujian algoritma C4.5 menggunakan metode *Cross Validation*, menunjukkan bahwa nilai *accuracy* sebesar 88,63% $\pm$ 0,83% dengan *micro average* sebesar 88,63%. Hasil ini menunjukkan bahwa model klasifikasi yang dibangun mampu mengklasifikasikan status perkara dengan tingkat ketepatan sebesar 88,63%.

Hasil evaluasi menunjukkan bahwa kelas Pembuatan Akta Cerai memperoleh nilai *precision* sebesar 89,38% dan *recall* sebesar 98,81%. Nilai tersebut menunjukkan bahwa model mampu mengidentifikasi sebagian besar data pada kelas tersebut dengan tingkat ketepatan yang sangat baik. Selain itu, kelas Pemberitahuan Putusan memperoleh *precision* sebesar 85,48% dan *recall* sebesar 98,15%, sehingga model juga mampu melakukan klasifikasi dengan baik pada kelas tersebut.

Berbeda dengan kedua kelas tersebut, kelas Minutasi memperoleh *precision* sebesar 65,67%, namun hanya memiliki *recall* sebesar 18,03%. Hal ini menunjukkan bahwa ketika model memprediksi suatu data sebagai Minutasi, sebagian besar prediksi tersebut benar. Akan tetapi, model masih belum mampu mengenali sebagian besar data yang sebenarnya termasuk ke dalam kelas Minutasi, sehingga masih banyak data yang diklasifikasikan ke kelas lain.

Sementara itu, beberapa kelas lainnya seperti Pemberitahuan Putusan Banding, Penetapan Ikrar Talak, Permohonan Kasasi, Pemberitahuan Hasil Putusan Verzet, Pemberitahuan Putus Kasasi, Penetapan Sidang Ikrar Talak, Putusan Perlawanan Verzet, Penyerahan Kontra Memori Kasasi, Pencabutan Perkara Banding, Pengiriman Berkas PK, Persidangan, Penunjukan Panitera Pengganti Ikrar Talak, Penetapan Majelis Hakim Ikrar Talak, Pengiriman Berkas Banding, dan Penetapan Jadwal Sidang Verzet memperoleh nilai *precision* dan *recall* sebesar 0%. Nilai tersebut menunjukkan bahwa model belum mampu mengenali kelas-kelas tersebut sehingga tidak menghasilkan prediksi pada kelas tersebut.

Kondisi tersebut dipengaruhi oleh ketidakseimbangan distribusi data (*class imbalance*). Berdasarkan *Confusion Matrix*, jumlah data pada kelas Pembuatan Akta Cerai mencapai 3.942 data, sedangkan sebagian besar kelas lainnya hanya memiliki 1 hingga 8 data. Perbedaan jumlah data yang sangat besar menyebabkan algoritma C4.5 lebih banyak mempelajari pola dari kelas mayoritas dibandingkan kelas minoritas. Akibatnya, model cenderung mengklasifikasikan data ke dalam kelas yang memiliki jumlah data lebih banyak sehingga kemampuan mengenali kelas minoritas menjadi rendah.

Hasil tersebut juga terlihat pada struktur *decision tree* yang dihasilkan. Pohon keputusan menunjukkan bahwa atribut Lama Proses menjadi *root node* karena memiliki nilai *information gain* tertinggi. Selanjutnya, atribut tersebut digunakan kembali pada beberapa cabang pohon keputusan bersama atribut Klasifikasi Perkara untuk menghasilkan keputusan akhir. Sebagian besar daun (*leaf node*) pada pohon keputusan menghasilkan kelas Pembuatan Akta Cerai, Pemberitahuan Putusan, dan Minutasi, sedangkan kelas lainnya tidak muncul sebagai hasil prediksi. Hal ini menunjukkan bahwa model lebih banyak membentuk aturan berdasarkan kelas yang memiliki jumlah data dominan.

Secara keseluruhan, algoritma C4.5 mampu menghasilkan model klasifikasi dengan tingkat akurasi yang cukup baik, yaitu 88,63%. Namun demikian, ketidakseimbangan distribusi data menyebabkan kemampuan model dalam mengenali kelas minoritas masih rendah. Oleh karena itu, pada penelitian selanjutnya dapat diterapkan teknik penyeimbangan data seperti SMOTE, *oversampling*, atau *undersampling* agar model mampu mempelajari pola dari seluruh kelas secara lebih merata sehingga nilai *precision* dan *recall* pada kelas minoritas dapat meningkat [11].

3.3 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma C4.5 mampu membentuk model klasifikasi status perkara berdasarkan data penyelesaian perkara di Pengadilan Agama dengan tingkat performa yang baik. Hal ini menunjukkan bahwa algoritma C4.5 dapat digunakan untuk mengidentifikasi pola hubungan antara atribut Klasifikasi Perkara dan Lama Proses terhadap Status Perkara melalui pembentukan aturan (*decision tree*) yang mudah dipahami.

Berdasarkan *decision tree* yang dihasilkan, atribut Lama Proses menjadi atribut yang paling berpengaruh dalam proses klasifikasi karena dipilih sebagai akar (*root*) pohon keputusan. Pemilihan atribut tersebut menunjukkan bahwa lama penyelesaian suatu perkara memiliki nilai informasi (*information gain*) tertinggi dibandingkan atribut lainnya, sehingga mampu membedakan status perkara secara lebih efektif. Selanjutnya, atribut Klasifikasi Perkara berperan sebagai atribut pendukung yang membantu memperjelas proses pengambilan keputusan pada beberapa cabang pohon keputusan.

Performa model yang diperoleh menunjukkan bahwa algoritma C4.5 telah mampu mengenali pola pada sebagian besar data, terutama pada kelas dengan jumlah data yang dominan. Sebaliknya, kemampuan model dalam mengklasifikasikan beberapa kelas dengan jumlah data yang relatif sedikit masih belum optimal. Kondisi ini dipengaruhi oleh distribusi data yang tidak seimbang (*imbalanced dataset*), sehingga proses pembelajaran model lebih banyak dipengaruhi oleh kelas mayoritas dibandingkan kelas minoritas.

Temuan tersebut sejalan dengan karakteristik algoritma C4.5 yang membangun aturan berdasarkan nilai *information gain* tertinggi sehingga cenderung menghasilkan pohon keputusan yang mengikuti distribusi data pelatihan. Oleh karena itu, apabila terdapat ketidakseimbangan jumlah data antar kelas, maka model akan lebih mudah mengenali kelas yang memiliki jumlah data lebih banyak dibandingkan kelas dengan jumlah data yang sedikit.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa algoritma C4.5 efektif digunakan untuk melakukan klasifikasi status perkara pada data Pengadilan Agama karena mampu menghasilkan model yang memiliki tingkat akurasi yang baik serta aturan klasifikasi yang mudah diinterpretasikan. Hasil tersebut diharapkan dapat menjadi informasi pendukung bagi pihak Pengadilan Agama dalam memahami pola penyelesaian perkara serta menjadi dasar bagi pengembangan penelitian selanjutnya menggunakan teknik penanganan data tidak seimbang atau algoritma klasifikasi lainnya untuk meningkatkan performa model pada seluruh kelas.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma C4.5 berhasil diterapkan untuk mengklasifikasikan status perkara perceraian pada data Sistem Informasi Penelusuran Perkara (SIPP) Pengadilan Agama Cikarang Tahun 2025. Penelitian menggunakan metodologi CRISP-DM yang meliputi tahapan Business Understanding, Data Understanding, Data Preparation, Modeling, dan Evaluation sehingga proses pembangunan model dilakukan secara sistematis. Dataset yang digunakan berjumlah 4.554 data perkara dengan atribut Klasifikasi Perkara dan Lama Proses sebagai prediktor serta Status Perkara sebagai label.

Hasil pemodelan menunjukkan bahwa atribut Lama Proses memiliki nilai information gain tertinggi sehingga dipilih sebagai akar (root node) pada pohon keputusan. Atribut tersebut menjadi faktor yang paling berpengaruh dalam menentukan status perkara, sedangkan atribut Klasifikasi Perkara berfungsi sebagai atribut pendukung pada beberapa cabang pohon keputusan. Model yang dihasilkan mampu membentuk aturan klasifikasi yang mudah dipahami dan dapat digunakan untuk menjelaskan hubungan antara karakteristik perkara dengan status perkara yang dihasilkan.

Berdasarkan hasil evaluasi menggunakan Cross Validation, model memperoleh nilai accuracy sebesar $88,63\% \pm 0,83\%$, yang menunjukkan bahwa algoritma C4.5 memiliki performa yang baik dalam mengklasifikasikan status perkara perceraian. Nilai precision dan recall yang tinggi pada kelas mayoritas menunjukkan bahwa model mampu mengenali pola data dengan baik. Namun, performa pada beberapa kelas minoritas masih rendah karena distribusi data yang tidak seimbang sehingga model lebih banyak mempelajari pola dari kelas yang memiliki jumlah data lebih besar.

Secara keseluruhan, penelitian ini berhasil menjawab tujuan penelitian, yaitu membangun model klasifikasi status perkara perceraian menggunakan algoritma C4.5 pada data SIPP Pengadilan Agama Cikarang Tahun 2025. Model yang dihasilkan dapat dimanfaatkan sebagai pendukung analisis data dan pengambilan keputusan karena mampu memberikan informasi mengenai pola penyelesaian perkara secara lebih cepat, objektif, dan mudah dipahami. Untuk penelitian selanjutnya, disarankan menerapkan teknik penyeimbangan data seperti SMOTE, *oversampling*, atau *undersampling* serta membandingkan algoritma C4.5 dengan algoritma klasifikasi lainnya agar kemampuan model dalam mengenali kelas minoritas dapat ditingkatkan dan menghasilkan performa klasifikasi yang lebih optimal.

REFERENCES

- [1] N. S. Manna, S. Doriza, and M. Oktaviani, "Cerai Gugat: Telaah Penyebab Perceraian Pada Keluarga di Indonesia," *J. AL-AZHAR Indones. SERI Hum.*, vol. 6, no. 1, pp. 11–21, 2021.
- [2] C. P. Agama, "Sistem Informasi Penelusuran Perkara (SIPP)." [Online]. Available: <https://sipp.pa-cikarang.go.id/>
- [3] F. Sholikhah and D. Kumalaeni, "Sistem informasi penelusuran perkara (sipp);," vol. 1, no. 1, pp. 38–46, 2017.
- [4] Z. Fatah and R. D. Saputra, "Algoritma Keputusan C4.5 Untuk Klasifikasi Data Medis Pasien," vol. 5, no. April, pp. 113–120, 2026.
- [5] A. I. Silitonga, D. A. Nasution, H. T. Naibaho, S. Apriani, and D. P. L. Bantu, "KLASIFIKASI FAKTOR PENYEBAB PERCERAIAN," vol. 11, no. 1, pp. 8–15, 2025.
- [6] D. Yusuf, S. Bahri, and A. Larasati, "DECISION TREE MENGGUNAKAN ALGORITMA C4 . 5 UNTUK," *J. Ilm. Teknol. Inf.*, vol. 2, no. 2, pp. 97–106, 2021.
- [7] C. Schröer, F. Kruse, and J. M. Gómez, "ScienceDirect ScienceDirect A Systematic Literature Review A Systematic Literature Review on Applying Process Model on Applying CRISP-DM Process Model," *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
- [8] T. Y. Pratama and Armansyah, "Decision Tree C4 . 5 dengan Teknik Information Gain Untuk Klasifikasi Pemilihan Program Studi Tingkat Lanjut," *J. Inf. Syst. Res.*, vol. 5, no. 4, pp. 1042–1052, 2024, doi: 10.47065/josh.v5i4.5643.
- [9] I. M. A. O. Gunawan, I. D. A. I. Saraswati, I. D. G. R. Agung, and I. P. E. Putra, "Klasifikasi Penyakit Jantung Menggunakan Algoritma Decision Tree Series C4 . 5 Dengan Rapidminer," *J. Teknol. Dan Sist. Inf. Bisnis*, vol. 5, no. 2, pp. 73–83, 2023.
- [10] S. P. Andani, R. F. Purnomo, and Y. Yuniarthe, "Penerapan Algoritma C4 . 5 untuk Klasifikasi Tingkat Kedisiplinan Siswa Sekolah Menengah," *J. Informatics, Electr. Electron. Eng.*, vol. 5, no. 1, pp. 40–46, 2025, doi: 10.47065/jieee.v5i1.2630.
- [11] W. Rahayu *et al.*, "Synthetic Minority Oversampling Technique (SMOTE) for Boosting the Accuracy of C4 . 5 Algorithm Model," *J. Artif. Intell. Eng. Appl.*, vol. 3, no. 3, pp. 625–230, 2024.