PERBANDINGAN ALGORITMA DECISION TREE DAN RANDOM FOREST DALAM KLASIFIKASI POLA DURASI PENYELESAIAN PERKARA PERCERAIAN BERBASIS DATA SIPP DI PENGADILAN NEGERI BITUNG

Constantine Kandowangko[1], Audy Aldrin Kenap[2], Quido Conferti Kainde[3]
[1,2,3] Teknik Informatika, Fakultas Teknik, Universitas Negeri Manado
e-mail: [1]constantinekandowangko@gmail.com, [2]audyakenap@unima.ac.id,
[3]quidokainde@unima.ac.id

ABSTRAK

Kasus perceraian merupakan salah satu jenis perkara perdata dengan volume tinggi yang memerlukan pengelolaan data tidak hanya secara administratif, tetapi juga analitis. Data dari sistem penelusuran perkara (SIPP) berpotensi digunakan untuk mengidentifikasi pola penyelesaian perkara secara lebih objektif. Penelitian ini bertujuan untuk mengklasifikasikan durasi penyelesaian kasus perceraian di Pengadilan Negeri Bitung ke dalam tiga kategori cepat, sedang, dan lama dengan menggunakan algoritma Decision Tree dan Random Forest. Kumpulan data mencakup 824 kasus dari periode 2021–2025 yang telah melalui tahap pembersihan, seleksi, pelabelan, dan transformasi fitur. Label durasi ditentukan menggunakan pendekatan kuartil, dengan ambang batas Q1 sebesar 37 hari dan Q3 sebesar 76 hari. Fitur-fitur yang digunakan meliputi jumlah hari persidangan, status putusan, serta bulan, hari dalam seminggu, dan kuartal pendaftaran perkara. Data dibagi dengan rasio 80:20 menggunakan metode stratified split dan selanjutnya dievaluasi berdasarkan akurasi, presisi, recall, skor F1, confusion matrix, dan validasi silang cross-validation. Hasil penelitian menunjukkan bahwa model Random Forest Tuned memiliki kinerja terbaik dengan akurasi 70,30%, presisi 0,7084, recall 0,7030, dan skor F1 0,6934 sehingga mengungguli model Decision Tree yang mencapai akurasi 59,39% dan skor F1 0,5954. Model dengan kinerja terbaik tersebut kemudian diimplementasikan ke dalam dasbor prediksi berbasis Streamlit. Hasil ini menunjukkan bahwa Random Forest Tuned menawarkan pendekatan yang lebih andal untuk mengklasifikasikan durasi kasus perceraian berdasarkan data administratif SIPP.

Kata Kunci: *Decision Tree, Klasifikasi, Perceraian, Random Forest, SIPP.*

ABSTRACT

Divorce cases are one of the high-volume types of civil cases that require data management not only from an administrative perspective but also from an analytical one. Data from the Case Tracking System (SIPP) can potentially be used to identify patterns in case resolution more objectively. This study aims to classify the duration of divorce case resolution at the Bitung District Court into three categories—fast, moderate, and slow—using the Decision Tree and Random Forest algorithms. The dataset comprises 824 cases from the 2021–2025 period that have undergone data cleaning, selection, labeling, and feature transformation. Duration labels were determined using a quartile

ISSN 2798-141X

approach, with a Q1 threshold of 37 days and a Q3 threshold of 76 days. The features used include the number of trial days, the status of the judgment, as well as the month, day of the week, and quarter of case filing. The data was split in an 80:20 ratio using the stratified split method and subsequently evaluated based on accuracy, precision, recall, F1 score, confusion matrix, and cross-validation. The results show that the Tuned Random Forest model performed best with an accuracy of 70.30%, precision of 0.7084, recall of 0.7030, and an F1 score of 0.6934, outperforming the Decision Tree model, which achieved an accuracy of 59.39% and an F1 score of 0.5954. The model with the best performance was then implemented into a Streamlit-based prediction dashboard. These results indicate that the Tuned Random Forest offers a more reliable approach for classifying the duration of divorce cases based on SIPP administrative data.

Keywords: *Classification, Decision Tree, Divorce, Random Forest, SIPP.*

PENDAHULUAN

Perceraian merupakan masalah sosial dan hukum yang trennya terus meningkat di Indonesia (Kurnia, 2021), oleh karena itu, penanganan kasus-kasus ini memerlukan dukungan berupa data pengadilan yang terorganisasi dan dapat diukur. Sistem Informasi Penelusuran Perkara (SIPP) menyimpan berbagai rincian administratif perkara, mulai dari tanggal pendaftaran hingga status putusan; namun, penggunaannya masih terutama berfokus pada input data dan pelaporan, dengan pemanfaatan yang terbatas untuk analisis berbasis data (Paramita dkk, 2023). Salah satu aspek yang berpotensi diperoleh dari data ini adalah durasi penyelesaian perkara, yang menjadi indikator awal efisiensi peradilan. Mengingat pola terkait durasi tidak selalu terlihat jelas dari data mentah, diperlukan pendekatan *data mining* dan *machine learning* untuk mengubah data administratif tersebut menjadi informasi yang lebih bernilai.

Pemanfaatan *machine learning* dalam sistem peradilan semakin signifikan seperti dalam memprediksi tingkat perceraian (Pasha, 2025) dan putusan perceraian berdasarkan data kasus (Rahmadini & Santoso, 2025), yang menunjukkan bahwa data peradilan dapat diolah untuk menghasilkan prediksi dan informasi guna mendukung pengambilan keputusan. Namun, sejumlah penelitian terdahulu mengenai data perceraian masih berfokus pada pendekatan *clustering* (Kurnia, 2021) meskipun metode ini mampu mengidentifikasi pola kelompok, hasilnya tidak menghasilkan model yang dapat secara langsung mengategorikan data baru ke dalam kategori tertentu. Oleh karena itu, penelitian ini menerapkan pendekatan klasifikasi menggunakan algoritma *decision tree* yang dipilih karena struktur aturannya mudah diinterpretasikan dan terbukti efektif untuk mengklasifikasikan data administratif (Hafiz dkk, 2026), serta metode *Random Forest*, yang menggabungkan beberapa *decision tree* melalui pendekatan *ensemble* untuk menghasilkan prediksi yang lebih tangguh (Tambuwun dkk, 2026) dan telah terbukti lebih unggul dibandingkan *decision tree* standar dalam berbagai skenario klasifikasi (Silvana dkk, 2025).

Berdasarkan pertimbangan tersebut, penelitian ini bertujuan untuk mengklasifikasikan waktu penyelesaian perkara perceraian di Pengadilan Negeri Bitung

ISSN 2798-141X

ke dalam kategori cepat, sedang, dan lama menggunakan algoritma *Decision Tree* dan *Random Forest*, sekaligus mengimplementasikan model dengan kinerja terbaik ke dalam dasbor prediksi berbasis *Streamlit* yang dirancang sebagai sarana pemantauan awal bagi pihak pengadilan.

KAJIAN TEORI

Data mining adalah proses pengolahan data yang digunakan untuk mengidentifikasi pola atau informasi tersembunyi, proses ini mencakup tahapan mulai dari pembersihan data, pemilihan fitur, dan pembangunan model hingga evaluasi hasil. Klasifikasi merupakan metode *data mining* di mana data ditempatkan ke dalam kelas-kelas tertentu berdasarkan pola yang diperoleh dari data berlabel (Helmud dkk, 2024), sehingga metode ini sesuai untuk penelitian yang bertujuan mengategorikan data ke dalam kelas pendek, sedang, dan panjang. Penggunaan kecerdasan buatan dalam sistem peradilan dapat meningkatkan efisiensi dan konsistensi, asalkan teknologi tersebut berfungsi sebagai alat bantu dan bukan sebagai pengganti diskresi manusia (Rodríguez-Salcedo dkk, 2025).

Pohon keputusan (*decision tree*) adalah algoritma klasifikasi yang membangun model dalam bentuk struktur pohon, dalam model ini, setiap simpul menguji suatu atribut, cabang menunjukkan hasil pengujian, dan simpul daun merepresentasikan kelas akhir. Keunggulan utamanya adalah kemudahan interpretasi, meskipun terdapat risiko *overfitting* jika struktur pohon menjadi terlalu dalam. *Random Forest* merupakan pengembangan dari pohon keputusan, metode ini menghasilkan sejumlah pohon dari berbagai subset data dan fitur, kemudian menentukan hasil akhir melalui proses pemungutan suara (*voting*). Akibatnya, metode ini cenderung lebih stabil, menangani variasi data secara lebih efektif, dan mengurangi risiko *overfitting* (Torondek dkk, 2025).

Evaluasi model bertujuan untuk menilai kemampuan algoritma dalam mengklasifikasikan data uji (Sumual dkk, 2025), dan mengandalkan metrik seperti akurasi, presisi, *recall*, serta skor F1. Evaluasi sebaiknya tidak hanya bergantung pada akurasi, karena metrik ini tidak selalu mencerminkan kinerja model pada seluruh kelas (Pinem & Putra, 2025). Selain itu, digunakan pula *Confusion Matrix*, karena matriks ini menampilkan distribusi prediksi yang benar dan salah pada setiap kategori (Wibawa dkk, 2025).

METODE PENELITIAN

Alat dan Bahan

Penelitian ini menggunakan laptop yang dilengkapi dengan prosesor AMD Ryzen 3, RAM 8 GB, dan SSD 256 GB. Perangkat lunak yang diinstal meliputi *Python*, Google *Colab*, Visual Studio Code, dan *Streamlit* (untuk dasbor), dengan dukungan pustaka *Pandas, NumPy, Scikit-learn, Matplotlib,* dan *Seaborn*. Materi penelitian terdiri dari data kasus perceraian yang bersumber dari sistem SIPP Pengadilan Negeri Bitung untuk periode 2021–2025, khususnya kasus-kasus yang telah memiliki putusan akhir berkekuatan hukum tetap. Setelah melalui tahap pembersihan dan seleksi data, diperoleh

ISSN 2798-141X

824 data yang valid, yang mencakup nomor perkara, tanggal pendaftaran, tanggal putusan, durasi perkara, jumlah hari sidang, dan status putusan. Variabel durasi perkara digunakan untuk menentukan kelas target, bukan sebagai fitur masukan, guna mencegah terjadinya kebocoran data (*data leakage*) pada model.

Jalannya Penelitian

1. Pengumpulan Data
   Data mengenai kasus perceraian untuk periode 2021–2025 diperoleh dari sistem SIPP Pengadilan Negeri Bitung, dengan batasan pemilihan hanya pada kasus yang telah memiliki putusan akhir. Pembersihan data mencakup penghapusan data kosong dan duplikat, perbaikan format tanggal, serta penyesuaian data dengan periode penelitian.
2. Pembersihan Data
   Pembersihan data dilakukan dengan menghapus data kosong, memperbaiki format tanggal, menghilangkan duplikat, dan memastikan data berada dalam rentang waktu penelitian. Data yang tidak memenuhi kriteria tersebut dikeluarkan dari proses pemodelan.
3. Pembentukan Kategori Klasifikasi
   Label untuk kategori durasi dibuat berdasarkan variabel durasi kasus menggunakan pendekatan kuartil, yaitu kuartil pertama (Q1) sebesar 37 hari ditetapkan sebagai ambang batas untuk kategori "Cepat", sedangkan kuartil ketiga (Q3) sebesar 76 hari ditetapkan sebagai ambang batas untuk kategori "Lama" seperti pada Tabel.1

   Tabel 1. Distribusi Kategori Durasi Penyelesaian Perkara

Kategori	**Threshold**	**Jumlah Perkara**
Cepat	≤ 37 hari	209
Sedang	37–76 hari	410
Lama	> 76 hari	205
Total		824

4. Rekayasa Fitur
   Fitur *input* yang digunakan mencakup jumlah sidang pengadilan, status putusan (yang dikodekan secara numerik melalui *label encoding*, serta bulan, hari dalam seminggu, dan kuartal pendaftaranyang semuanya diturunkan dari tanggal pendaftaran pada Tabel 2.

   Tabel 2. Fitur Input yang Digunakan dalam Pemodelan

No.	Nama Fitur	Keterangan
1	Jumlah Sidang	Total sidang yang dilaksanakan dalam satu perkara
2	Status Putusan	Status putusan akhir yang telah dikodekan secara numerik
3	bulan_register	Bulan pendaftaran perkara
4	hari_dalam_minggu_register	Hari dalam minggu saat perkara didaftarkan
5	kuartal_register	Kuartal pendaftaran perkara

ISSN 2798-141X

5. Pembagian Data
   Dataset dibagi menggunakan metode *stratified split* dengan rasio 80:20, yang menghasilkan 659 sampel pelatihan dan 165 sampel pengujian, sehingga memastikan distribusi kelas pada kedua *subset* tetap representatif terhadap keseluruhan *dataset*.
6. Pemodelan dan Evaluasi
   Model dikembangkan menggunakan *DecisionTreeClassifier* dan *RandomForestClassifier* untuk model *Random Forest*, hiperparameter *n_estimators,max_depth,min_samples_split,* dan *min_samples_leaf)* dioptimalkan menggunakan *GridSearchCV* dan validasi silang 5-lipat *(5-fold cross-validation).* Kinerja model dievaluasi berdasarkan akurasi, presisi, *recall*, skor F1, matriks konfusi, dan validasi silang. selanjutnya, model dengan kinerja terbaik diimplementasikan dalam dasbor berbasis *Streamlit*..

Analisis Data

Analisis data dilakukan dengan membandingkan kinerja model *Decision Tree* dan *Random Forest* pada data uji, serta mengevaluasi akurasi, presisi, *recall*, dan skor F1 menggunakan persamaan 1 hingga 4.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP}\ ) \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

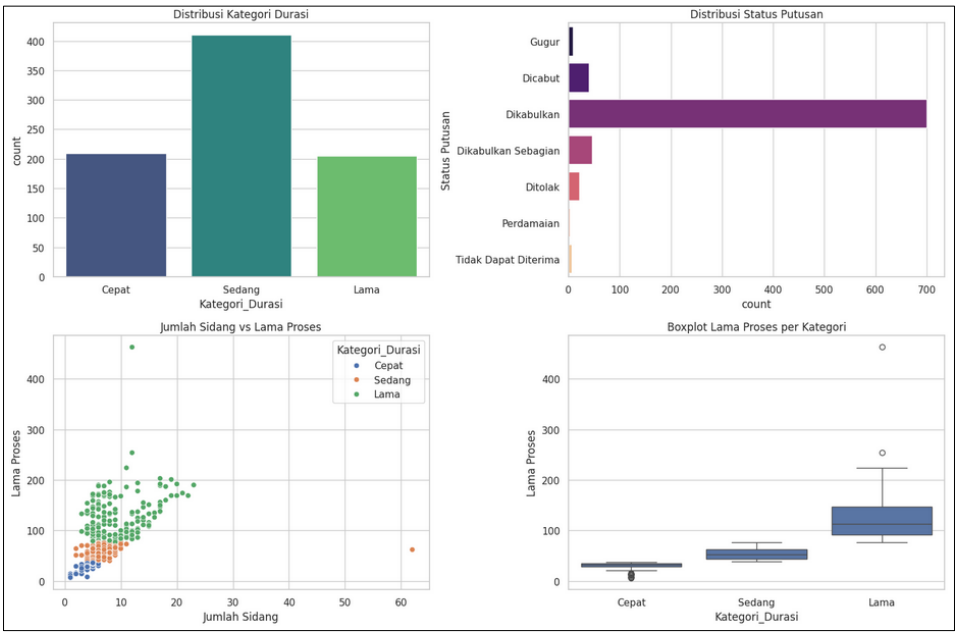
$$F1\text{-}Score = \frac{2\cdot(Precision\cdot Recall)}{Precision+Recall} \quad (4)$$

Keterangan: *TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative*. Karena penelitian ini melibatkan analisis multi-kelas, hasil evaluasi juga dinilai menggunakan metode *macro-averaging* dan *weighted averaging*.

HASIL DAN PEMBAHASAN

Setelah tahap pembersihan dan seleksi data, sebanyak 824 catatan yang valid disertakan dalam penelitian ini, dengan rata-rata durasi kasus sebesar 65,21 hari (berkisar antara minimal 6 hari hingga maksimal 463 hari) dan rata-rata 6,67 hari sidang per kasus. Sebagaimana ditunjukkan pada Gambar 1, status putusan dikabulkan merupakan kategori yang paling sering ditemukan, mencakup 700 kasus.

ISSN 2798-141X



Gambar 1. Distribusi Kategori Durasi, Status Putusan, dan Hubungan Jumlah Sidang terhadap Lama Proses

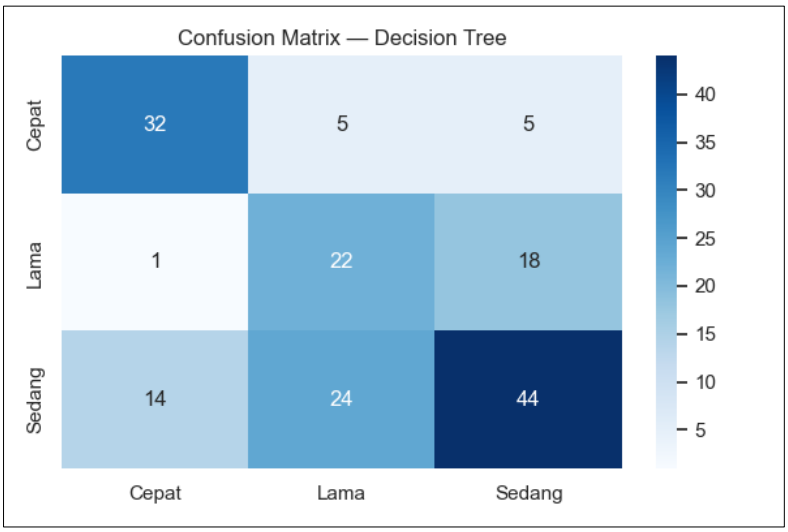
Hasil klasifikasi menunjukkan bahwa kategori Sedang merupakan kelompok terbesar (410 kasus; 49,76%), diikuti oleh kategori "Cepat" (209; 25,36%) dan Lambat (205; 24,88%), yang mengindikasikan bahwa sebagian besar kasus memiliki durasi sedang.

Model pohon keputusan, yang dilatih menggunakan 659 *dataset* dan diuji pada 165 *dataset*, mencapai tingkat akurasi sebesar 59,39% pada Tabel 3. Kelas Cepat merupakan kelas yang paling mudah diidentifikasi (F1 = 0,72), sedangkan kelas Lambat menghasilkan skor F1 terendah (0,48). Hal ini menunjukkan bahwa model masih mengalami kesulitan dalam membedakan kasus pada kategori Lambat sebagaimana terlihat pada Gambar 2, kesalahan prediksi sering terjadi antara kategori Sedang dan Lambat akibat terbatasnya ketersediaan fitur administratif..

Tabel 3. Hasil Evaluasi Model Decision Tree

Kelas	Precision	Recall	F1-Score	Support
Cepat	0,68	0,76	0,72	42
Lama	0,43	0,54	0,48	41
Sedang	0,66	0,54	0,59	82
Accuracy	-	-	0,59	165
Macro Avg	0,59	0,61	0,60	165
Weighted Avg	0,61	0,59	0,60	165

ISSN 2798-141X



Gambar 2. Confusion Matrix Model Decision Tree

Model *Random Forest* dioptimalkan menggunakan *GridSearchCV* dengan *5-fold cross-validation*, yang menghasilkan parameter optimal *n_estimators = 100, max_depth = 10, min_samples_split = 2*, dan *min_samples_leaf = 4*, dengan skor F1 *cross-validation* terbaik sebesar 0,7201, sebagaimana ditunjukkan pada hasil di Tabel 4..

Tabel 4. Hyperparameter Terbaik Random Forest

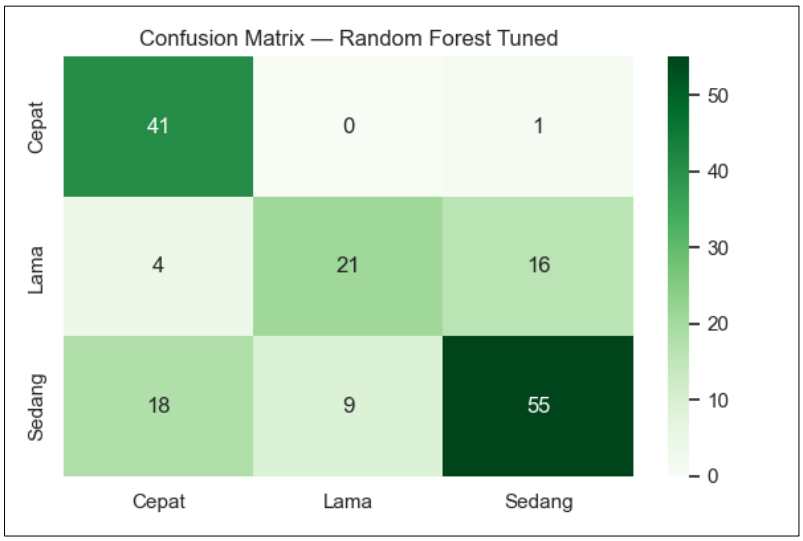
Parameter	Nilai Terbaik
n_estimators	100
max_depth	10
min_samples_split	2
min_samples_leaf	4
Best CV F1-Score	0,7201

Random Forest Tuned mencapai kinerja yang lebih baik, dengan akurasi 70,30%, presisi 0,7084, *recall* 0,7030, dan *F1-score* 0,6934 seperti pada Tabel 5. Kelas Cepat mencapai *recall* tertinggi 0,98 sedangkan kelas Lama tetap menjadi kategori yang paling sulit F1 = 0,56, meskipun hasilnya membaik dibandingkan dengan *decision tree* pada Gambar 3.

Tabel 5. Hasil Evaluasi Model Random Forest Tuned

Kelas	Precision	Recall	F1-Score	Support
Cepat	0,66	0,98	0,79	42
Lama	0,67	0,49	0,56	41
Sedang	0,75	0,67	0,71	82
Accuracy	-	-	0,70	165
Macro Avg	0,69	0,71	0,69	165
Weighted Avg	0,71	0,70	0,69	165

ISSN 2798-141X



Gambar 3. Confusion Matrix Model Random Forest Tuned

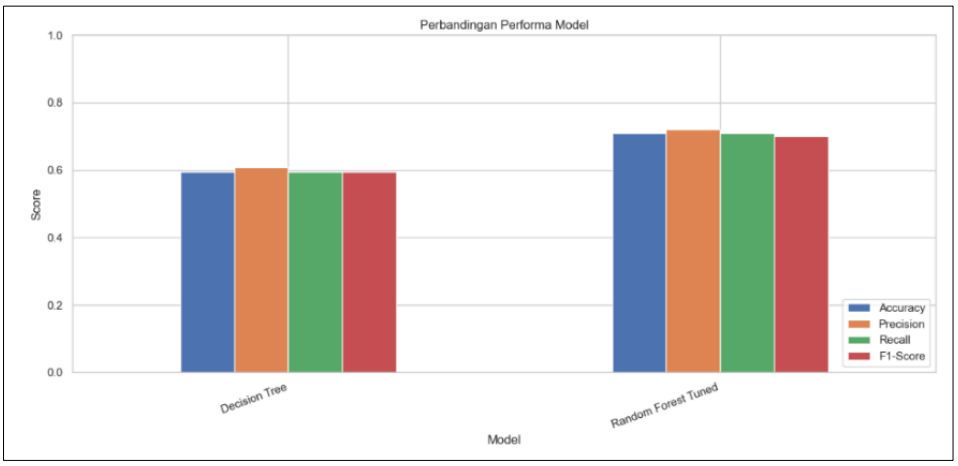
Perbandingan ketiga model menunjukkan bahwa *Random Forest Tuned* mengungguli model lainnya pada semua metrik; model ini melampaui *Decision Tree* sebesar 10,91 poin persentase dalam hal akurasi dan 0,0980 dalam skor F1 seperti pada perbandingan kinerja model dalam Tabel 6. dan grafik pada Gambar 4. Hasil ini sejalan dengan temuan bahwa model *Random Forest* cenderung lebih stabil dibandingkan *Decision Tree* dalam berbagai skenario klasifikasi (Silvana dkk, 2025) sebuah karakteristik yang dikaitkan dengan kemampuan pendekatan *ensemble* dalam memitigasi *overfitting* melalui *bootstrap sampling* dan pemungutan suara mayoritas *majority voting*. Meskipun demikian, tingkat akurasi sebesar 70,30% menunjukkan adanya ruang untuk peningkatan lebih lanjut, terutama mengingat penelitian ini hanya mengandalkan fitur administratif dan tidak memperhitungkan aspek lain seperti penjadwalan ulang negosiasi, kehadiran para pihak, atau hasil mediasi.

Tabel 6. Perbandingan Performa Model Klasifikasi

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	59,39%	0,6069	0,5939	0,5954
Random Forest Baseline	66,67%	0,6641	0,6667	0,6617
Random Forest Tuned	70,30%	0,7084	0,7030	0,6934

Analisis mengenai tingkat kepentingan fitur menunjukkan bahwa jumlah hari sidang merupakan fitur yang paling dominan 0,7315, diikuti oleh hari pendaftaran 0,0966, bulan pendaftaran 0,0954, status putusan 0,0392, dan kuartal register 0,0373 seperti pada Tabel 7 dan grafik bar pada Gambar 5. Hasil ini mengindikasikan bahwa intensitas sidang merupakan indikator administratif terkuat untuk membedakan kategori durasi, sedangkan fitur temporal dan status putusan memberikan kontribusi yang relatif kecil.

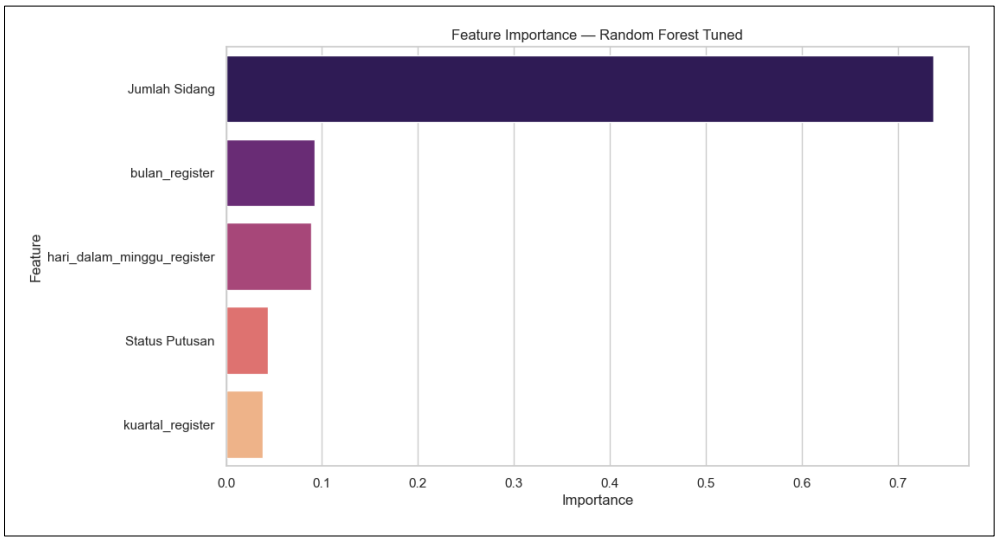
ISSN 2798-141X



Gambar 4. Diagram Batang Perbandingan Performa Model Klasifikasi

Tabel 7. Feature Importance Model Random Forest Tuned

Fitur	Nilai Importance
Jumlah Sidang	0,7315
hari_dalam_minggu_register	0,0966
bulan_register	0,0954
Status Putusan	0,0392
kuartal_register	0,0373

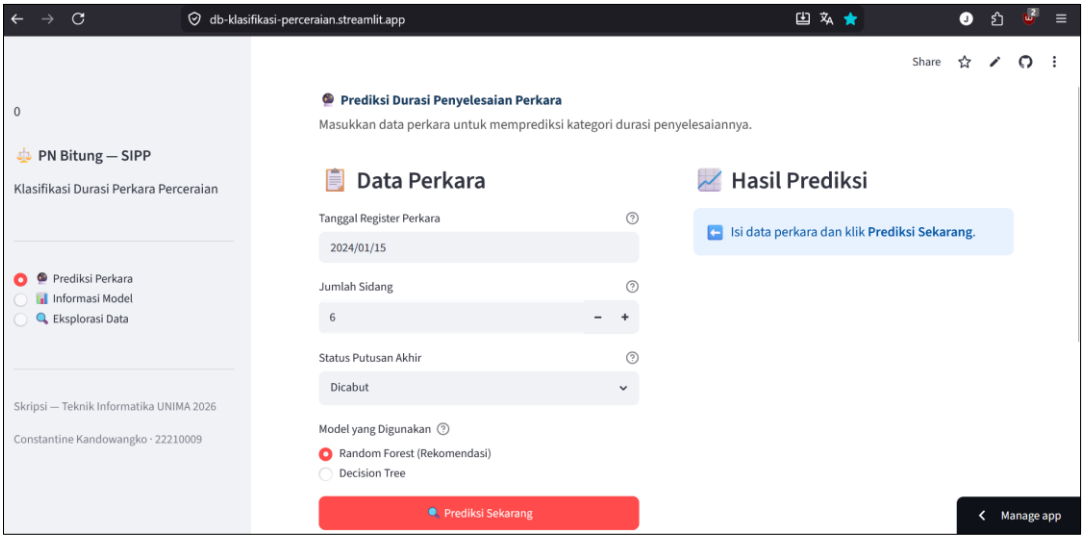


Gambar 5. Feature Importance Model Random Forest Tuned

Model dengan kinerja terbaik diimplementasikan dalam dasbor berbasis Streamlit yang mencakup halaman Prediksi Kasus, Informasi Model, dan Eksplorasi Data seperti yang ditujukan pada Gambar 6. Uji fungsional pada tiga skenario menunjukkan bahwa dasbor tersebut mampu memprediksi kategori Cepat, Sedang, dan Lama secara konsisten

ISSN 2798-141X

dengan tingkat keyakinan di atas 84% pada hasil pengujian Tabel 8. hasil ini selaras dengan hubungan antara jumlah tanggal sidang dan kategori durasi yang teridentifikasi dalam analisis kepentingan fitur.



Gambar 6. Tampilan Halaman Prediksi Perkara pada Dashboard

Tabel 8. Hasil Uji Fungsional Tiga Skenario Prediksi Dashboard

Skenario	Jumlah Sidang	Status Putusan	Tanggal Register	Prediksi	Confidence
Sidang Sedikit	3	Dikabulkan	2024-01-10	Cepat	84,0%
Sidang Sedang	7	Dikabulkan	2023-06-15	Sedang	88,6%
Sidang Banyak	15	Dikabulkan	2022-03-01	Lama	90,2%

Secara umum, model *Random Forest Tuned* lebih tepat digunakan dibandingkan *Decision Tree* untuk mengklasifikasikan durasi prosedur, karena pendekatan *ensemble* menghasilkan prediksi yang lebih stabil. Namun, model ini sebaiknya dipandang sebagai alat analisis, bukan sebagai dasar mutlak untuk pengambilan keputusan. oleh karena itu, hasil prediksi tetap harus ditafsirkan dengan mempertimbangkan realitas administratif serta kondisi spesifik dari setiap kasus.

KESIMPULAN

Penelitian ini berhasil menerapkan algoritma *Decision Tree* dan *Random Forest* untuk mengklasifikasikan durasi penyelesaian perkara perceraian di Pengadilan Negeri Bitung, berdasarkan 824 data SIPP dari periode 2021–2025, durasi tersebut dikategorikan sebagai cepat, sedang, atau lama menggunakan pendekatan berbasis kuartil. Model *Random Forest Tuned* menunjukkan kinerja terbaik dengan akurasi 70,30%, presisi

ISSN 2798-141X

0,7084, *recall* 0,7030, dan skor F1 0,6934 mengungguli *Decision Tree* hanya mencapai akurasi 59,39% dan skor F1 0,5954. Variabel jumlah hari sidang terbukti menjadi faktor paling berpengaruh dalam model, yang menunjukkan adanya korelasi kuat antara intensitas persidangan dan kategori durasi penyelesaian perkara. Model dengan kinerja terbaik tersebut diimplementasikan dalam dasbor berbasis *Streamlit* yang dirancang sebagai sarana pemantauan awal. Untuk penelitian selanjutnya, disarankan untuk menyertakan fitur yang lebih rinci seperti alasan perceraian, agenda sidang, hasil mediasi, kehadiran para pihak, dan kelengkapan berkas agar model dapat menangkap pola penyelesaian perkara perceraian secara lebih komprehensif.

DAFTAR PUSTAKA

Hafiz, M. F., Iskandar, A., Harahap, D. R., & Purba, S. S. T. (2026). Efektifitas Penerapan Algoritma Decision Tree pada Sistem E-Persuratan dalam Meningkatkan Kecepatan dan Akurasi Pengelolaan Dokumen. Jurnal Komputer Teknologi Informasi Sistem Komputer (JUKTISI), 5(1), 212-218.

Helmud, E., Fitriyani, F., & Romadiana, P. (2024). Classification comparison performance of supervised machine learning random forest and decision tree algorithms using confusion matrix. Jurnal Sisfokom (Sistem Informasi dan Komputer), 13(1), 92-97.

Kurnia, D. (2021). Pembangunan Model Clustering Dalam Pengelompokan Pengadilan Agama Berdasarkan Kasus Perceraian Dengan Menggunakan Algoritma K-Means. KOPERTIP: Jurnal Ilmiah Manajemen Informatika dan Komputer.

Paramita, C., & Rafrastara, F. A. (2023). Pemanfaatan Algoritma K-Means untuk Membuktikan Implementasi Undang-Undang Pelanggaran Hukum Korupsi di Pengadilan Negeri Banjarmasin. Jurnal Informatika: Jurnal Pengembangan IT, 8(2), 149-154.

Pasha, K. (2025). Prediksi Angka Perceraian Menggunakan Machine Learning. Buffer Informatika, 11(1), 1-6.

Pinem, T. H., & Putra, Z. P. (2025). Evaluasi kinerja algoritma klasifikasi deep learning dalam prediksi diabetes. Jurnal Ilmiah FIFO, 17(1), 17-28.

Rahmadini, R., & Santoso, B. J. (2025). Machine learning-based prediction of divorce verdicts using posita data and imbalanced data handling: A case study in Padang Sidempuan. International Journal of Advances in Data and Information Systems, 6(2), 460-478.

Rodríguez-Salcedo, E., Martínez-Bonilla, C., Pérez-Mayorga, B., Salame-Ortiz, M., Armas-Freire, P., Espín-Miniguano, A., & Pino-Loza, E. (2025). Evaluating AI decision tools in Ecuador's courts: efficiency, consistency, and uncertainty in legal judgments. Frontiers in Artificial Intelligence, 8, 1688209.

Silvana, R. A., Anggiani, N., Labib, A., Pratiwi, R. L., & Widanegsih, E. (2025). Perbandingan Kinerja Algoritma Decision Tree dan Random Forest dalam Memprediksi Kepuasan Penumpang Maskapai. Jurnal Ilmiah Sistem Informasi dan Ilmu Komputer, 5(3).

ISSN 2798-141X

Sumual, H., Sumual, H., & Ratumbuisang, K. F. (2025). Student Evaluation of Lecturer Performance Faculty of Engineering, Universitas Negeri Manado. International Journal of Information Technology and Education, 5(1), 24-39.

Tambuwun, F. B., Kenap, A. A., & Tangkawarow, I. R. H. T. (2026). Predicting ZNERGY Athletes' Swimming Times Using Random Forest Based on Lifestyle Physiological Exercise. Journal La Multiapp, 7(2), 500-510.

Torondek, D. C. H., Kaparang, D. R., & Pardanus, R. H. W. (2025). Implementasi Algoritma Random Forest untuk Klasifikasi Judul Skripsi berdasarkan Bidang Konsentrasi pada Jurusan PTIK Universitas Negeri Manado. Edutik: Jurnal Pendidikan Teknologi Informasi dan Komunikasi, 5(6), 1922-1935.

Wibawa, A. D., Suryani, S. D., & Pratasik, S. (2025). Classifying Electroencephalogram (EEG) Signals via Brain Activity Mapping to Distinguish Identified vs Unidentified Information. CommIT (Communication and Information Technology) Journal, 19(1), 101-114.

ISSN 2798-141X