

PERBANDINGAN PERFORMA DECISION TREE DAN RANDOM FOREST UNTUK KLASIFIKASI KONDISI BAYI PREMATUR BERBASIS WEBSITE

Nur Fauziah¹, Siska Atmawan Oktavia², Yunanri.W³, Dinolah⁴

fauziah120404@gmail.com¹, siska.atmawan.oktavia@uts.ac.id², yunanri.w@uts.ac.id³,
dinola@uts.ac.id⁴

Universitas Teknologi Sumbawa

ABSTRAK

Bayi prematur memiliki risiko komplikasi kesehatan yang jauh lebih tinggi dibandingkan bayi cukup bulan, mulai dari gangguan pernapasan hingga peningkatan risiko kematian neonatal. World Health Organization (WHO) mencatat sekitar 15 juta kelahiran prematur terjadi setiap tahun di dunia, sementara di Indonesia prevalensinya diperkirakan mencapai sekitar 10%, dengan Bayi Berat Lahir Rendah (BBLR) menjadi penyebab kematian neonatus terbesar kedua. Rumah Sakit H.L. Manambai Abdulkadir memiliki 200 data rekam medis bayi prematur yang selama ini hanya dimanfaatkan sebagai arsip dan belum diolah secara optimal untuk mendukung pengambilan keputusan klinis. Penelitian ini bertujuan menerapkan dan membandingkan performa algoritma Decision Tree dan Random Forest dalam mengklasifikasikan kondisi kesehatan bayi prematur menggunakan Orange Data Mining, serta mengimplementasikan model terbaik ke dalam sistem berbasis website. Penelitian menggunakan metode CRISP-DM terhadap 200 data rekam medis dengan tujuh atribut klinis. Hasil evaluasi menggunakan 10-fold cross validation menunjukkan algoritma Decision Tree memperoleh accuracy 0,980, precision 0,981, recall 0,980, dan F1-score 0,980, sedangkan Random Forest memperoleh accuracy, precision, recall, dan F1-score masing-masing sebesar 0,985. Berdasarkan hasil tersebut, Random Forest dipilih sebagai model terbaik dan diimplementasikan ke dalam sistem berbasis website menggunakan Python dan Streamlit. Hasil pengujian Black Box Testing menunjukkan seluruh fitur sistem berfungsi dengan baik dan dinyatakan valid, sehingga sistem ini berpotensi digunakan sebagai alat bantu klasifikasi kondisi kesehatan bayi prematur secara lebih cepat, objektif, dan terkomputerisasi.

Kata Kunci: Decision Tree, Random Forest, Klasifikasi, Bayi Prematur, Orange Data Mining, Crisp-Dm.

ABSTRACT

Premature infants face a significantly higher risk of health complications than full-term infants, ranging from respiratory distress to an increased risk of neonatal mortality. The World Health Organization (WHO) reports approximately 15 million preterm births occur worldwide each year, while in Indonesia the prevalence is estimated at around 10%, with Low Birth Weight (LBW) being the second leading cause of neonatal death. H.L. Manambai Abdulkadir Hospital has 200 medical records of premature infants that have largely been used only as archives and have not been optimally processed to support clinical decision-making. This study aims to apply and compare the performance of the Decision Tree and Random Forest algorithms in classifying the health condition of premature infants using Orange Data Mining, and to implement the best-performing model into a web-based system. The study employs the CRISP-DM method on 200 medical records with seven clinical attributes. The results of 10-fold cross validation show that the Decision Tree algorithm achieved an accuracy of 0.980, precision of 0.981, recall of 0.980, and F1-score of 0.980, while Random Forest achieved an accuracy, precision, recall, and F1-score of 0.985 each. Based on these results, Random Forest was selected as the best-performing model and implemented into a web-based system using Python and Streamlit. Black-box testing results show that all system features function properly and are declared valid, indicating that the system can be used as a tool to classify the health condition of premature infants more quickly, objectively, and in a computerized manner.

Keywords: Decision Tree, Random Forest, Classification, Premature Infant, Orange Data Mining, Crisp-Dm.

PENDAHULUAN

Perkembangan teknologi informasi saat ini telah mendorong pemanfaatan data di berbagai bidang kehidupan, termasuk bidang kesehatan. Rumah sakit sebagai fasilitas pelayanan kesehatan menghasilkan data rekam medis dalam jumlah besar yang mencakup berbagai informasi klinis pasien. Data tersebut tidak hanya berfungsi sebagai arsip, tetapi juga memiliki potensi besar untuk diolah menjadi informasi yang bernilai melalui pendekatan data mining, sehingga dapat mempermudah pelaksanaan analisis dan menjadi dasar dalam proses pengambilan keputusan berbasis data secara efektif dan efisien [1].

Salah satu permasalahan kesehatan yang masih menjadi perhatian serius di tingkat global adalah tingginya angka kelahiran bayi prematur. Menurut World Health Organization (WHO), bayi prematur adalah bayi yang lahir sebelum usia kehamilan mencapai 37 minggu. WHO mencatat sekitar 15 juta bayi prematur lahir setiap tahun di seluruh dunia, dan kondisi tersebut menjadi salah satu penyebab utama kematian bayi baru lahir [2]. Dibandingkan bayi yang lahir cukup bulan, bayi prematur lebih rentan mengalami berbagai komplikasi kesehatan, antara lain gangguan pernapasan, infeksi neonatal, gangguan pertumbuhan, serta peningkatan risiko kematian pada periode neonatal.

Di Indonesia, kelahiran prematur masih menjadi tantangan kesehatan yang cukup serius. Berdasarkan data Survei Status Gizi Indonesia (SSGI), prevalensi Bayi Berat Lahir Rendah (BBLR) di Indonesia tercatat sebesar 6,0%, sementara estimasi WHO dan UNICEF menyebutkan prevalensi kelahiran prematur di Indonesia berada di angka sekitar 10%. Selain itu, data Maternal Perinatal Death Notification (MPDN) Kementerian Kesehatan RI mencatat jumlah kematian neonatus pada tahun 2022 sebanyak 22.882 kasus, dengan BBLR sebagai penyebab langsung kematian terbesar kedua sebanyak 25% [3]. Kondisi tersebut menunjukkan bahwa bayi prematur masih memerlukan penanganan yang optimal, baik dari aspek medis maupun pemanfaatan data kesehatan untuk mendukung pengambilan keputusan berbasis data mining.

Rumah Sakit H.L. Manambai Abdulkadir merupakan salah satu rumah sakit yang menangani kasus bayi prematur di Kabupaten Sumbawa. Berdasarkan hasil observasi di ruang NICU, tersedia sebanyak 200 data rekam medis bayi prematur periode Agustus 2025 hingga Februari 2026 yang memuat berbagai atribut klinis, seperti jenis kelamin, berat badan lahir, usia gestasi, skor APGAR, saturasi oksigen, kondisi asfiksia, kondisi infeksi, serta status kesehatan bayi. Data tersebut memiliki potensi besar untuk dianalisis menggunakan teknik data mining guna menemukan pola kondisi kesehatan bayi prematur berdasarkan data historis pasien, namun selama ini masih lebih banyak dimanfaatkan sebagai arsip dan belum digunakan secara optimal. Dalam penelitian ini, proses klasifikasi dilakukan menggunakan perangkat Orange Data Mining karena memiliki antarmuka visual yang memudahkan proses pengolahan dan evaluasi data tanpa memerlukan penulisan kode program secara kompleks.

Salah satu metode dalam data mining yang dapat diterapkan untuk mengolah data kesehatan adalah klasifikasi. Sejumlah penelitian terdahulu menunjukkan bahwa algoritma Decision Tree dan Random Forest memiliki performa yang baik dalam menyelesaikan permasalahan klasifikasi pada berbagai bidang. Penelitian [4] menerapkan algoritma Decision Tree untuk prediksi kelahiran bayi prematur dengan hasil yang baik, namun hanya menggunakan satu algoritma sehingga belum diketahui perbandingan performanya dengan algoritma lain yang sejenis. Penelitian [5] membandingkan Decision Tree dan Random Forest untuk klasifikasi nutrisi pada makanan cepat saji dan menunjukkan bahwa Random Forest mengungguli Decision Tree. Penelitian [6] membandingkan kedua algoritma untuk memprediksi keberhasilan immunotherapy dan menemukan bahwa Random Forest menghasilkan prediksi yang lebih tepat. Penelitian [7] menerapkan kedua

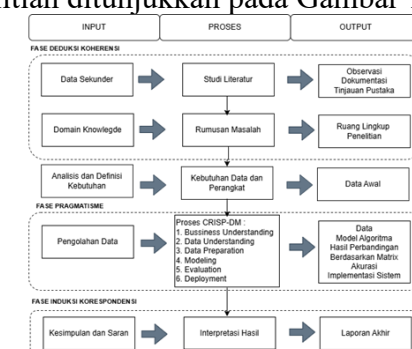
algoritma pada klasifikasi penerima bantuan sosial BPNT dan memperoleh tingkat akurasi yang tinggi dan relatif setara untuk kedua algoritma. Penelitian [8] membandingkan Decision Tree dan Random Forest untuk deteksi dini kesehatan mental mahasiswa dan menemukan bahwa Decision Tree sedikit unggul pada F1-score, meskipun Random Forest lebih stabil dan tahan terhadap overfitting. Penelitian [9] menunjukkan bahwa Random Forest lebih stabil dibandingkan Decision Tree pada klasifikasi gangguan tidur. Penelitian [10] menerapkan Decision Tree dan Random Forest untuk prediksi perdarahan pascasalin dan menunjukkan Random Forest memiliki nilai AUC yang lebih tinggi. Penelitian [11] membandingkan beberapa model machine learning untuk klasifikasi penyakit jantung dengan implementasi berbasis website dan menemukan bahwa Random Forest memiliki performa terbaik.

Berdasarkan tinjauan pustaka tersebut, sebagian besar penelitian terdahulu menerapkan algoritma Decision Tree dan Random Forest pada berbagai objek penelitian, seperti nutrisi makanan, immunotherapy, bantuan sosial, kesehatan mental, gangguan tidur, perdarahan pascasalin, dan penyakit jantung, namun belum ditemukan penelitian yang secara khusus membandingkan performa Decision Tree dan Random Forest pada klasifikasi kondisi kesehatan bayi prematur menggunakan data rekam medis di Rumah Sakit H.L. Manambai Abdulkadir. Penelitian [4] memang menggunakan objek bayi prematur, namun hanya menerapkan satu algoritma tanpa melakukan analisis perbandingan performa. Kekosongan tersebut menjadi research gap yang mendasari dilakukannya penelitian ini.

Berdasarkan uraian tersebut, penelitian ini bertujuan menerapkan dan membandingkan performa algoritma Decision Tree dan Random Forest dalam mengklasifikasikan kondisi kesehatan bayi prematur menggunakan Orange Data Mining berdasarkan metrik accuracy, precision, recall, dan F1-score, guna menentukan algoritma dengan performa terbaik. Model dengan performa terbaik selanjutnya diimplementasikan ke dalam sistem berbasis website menggunakan Python dan framework Streamlit, sehingga dapat digunakan sebagai alat bantu bagi tenaga kesehatan dalam melakukan klasifikasi kondisi kesehatan bayi prematur secara lebih cepat, objektif, dan terkomputerisasi di Rumah Sakit H.L. Manambai Abdulkadir.

METODOLOGI

Penelitian ini menggunakan pendekatan Cross Industry Standard Process for Data Mining (CRISP-DM) [13] sebagai kerangka kerja dalam proses perbandingan performa algoritma Decision Tree dan Random Forest untuk klasifikasi kesehatan bayi prematur. Tahapan penelitian terdiri atas Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Pada penelitian ini, tahap deployment diwujudkan dalam bentuk implementasi algoritma terbaik ke dalam sistem berbasis website. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

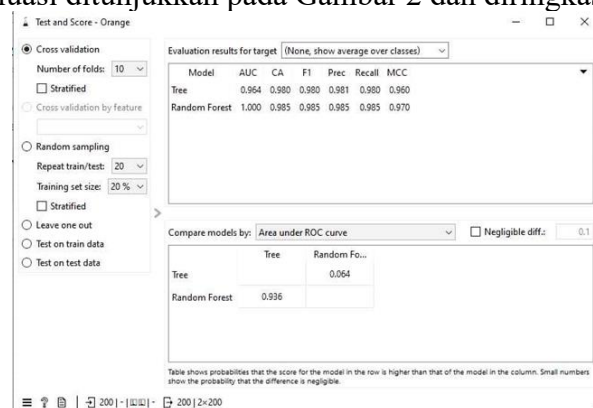
HASIL DAN PEMBAHASAN

A. Hasil Pengolahan Data

Pengolahan data pada penelitian ini dilakukan menggunakan Orange Data Mining [14] dengan metode CRISP-DM yang terdiri atas enam tahapan, yaitu business understanding, data understanding, data preparation, modeling, evaluation, dan deployment. Workflow yang dibangun terdiri atas tujuh widget yang saling terhubung, yaitu widget File untuk memuat dataset, widget Data Table untuk menampilkan data mentah, widget Preprocess untuk pra-pemrosesan data, widget Decision Tree dan Random Forest untuk membangun model klasifikasi, widget Test and Score untuk menguji dan mengevaluasi performa model, serta widget Confusion Matrix untuk menganalisis hasil prediksi setiap kelas secara lebih rinci.

B. Hasil Evaluasi Model

Evaluasi dilakukan menggunakan metode 10-fold cross validation pada widget Test and Score. Hasil evaluasi ditunjukkan pada Gambar 2 dan diringkas pada Tabel 2.

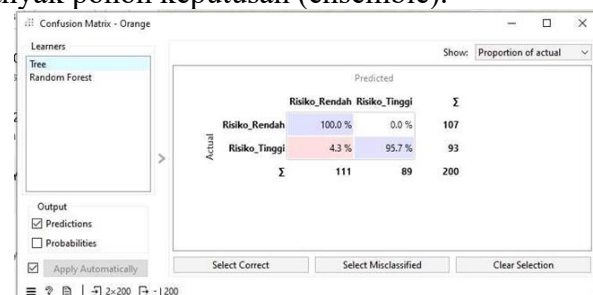


Gambar 2. Hasil Pengujian Widget Test and Score

Tabel 1. Hasil Evaluasi Model Decision Tree dan Random Forest

Algoritma	Accuracy	Precision	Recall	F1-Score
Decision Tree	0,980	0,981	0,980	0,980
Random Forest	0,985	0,985	0,985	0,985

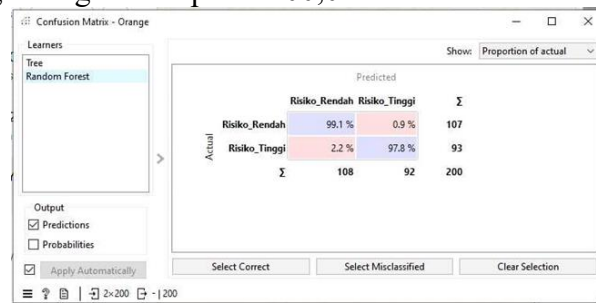
Berdasarkan Tabel 1, algoritma Random Forest memperoleh nilai accuracy, precision, recall, dan F1-score yang lebih tinggi (0,985) dibandingkan Decision Tree (accuracy 0,980; precision 0,981; recall 0,980; F1-score 0,980). Meskipun selisih nilai evaluasi kedua algoritma relatif kecil, hasil ini konsisten dengan temuan penelitian sebelumnya [5], [6], [10], [11] yang menunjukkan bahwa Random Forest cenderung memiliki performa yang lebih stabil dibandingkan Decision Tree karena menggabungkan hasil prediksi dari banyak pohon keputusan (ensemble).



Gambar 3. Confusion Matrix Decision Tree

Gambar 3 menunjukkan hasil confusion matrix Decision Tree. Model ini mampu mengklasifikasikan kelas Risiko_Tinggi dengan tingkat ketepatan 95,7% dan kelas

Risiko_Rendah dengan tingkat ketepatan 100,0%.



Gambar 4. Confusion Matrix Random Forest

Gambar 4 menunjukkan hasil confusion matrix Random Forest. Model ini mampu mengklasifikasikan kelas Risiko_Tinggi dengan tingkat ketepatan 97,8% dan kelas Risiko_Rendah dengan tingkat ketepatan 99,1%. Kesalahan klasifikasi yang terjadi relatif kecil dibandingkan jumlah data yang berhasil diprediksi dengan benar, sehingga menunjukkan bahwa Random Forest memiliki kemampuan yang baik dalam mengenali pola pada data bayi prematur.

C. Verifikasi Hasil Evaluasi Menggunakan Confusion Matrix

Untuk membuktikan validitas hasil pengolahan data pada Orange Data Mining, dilakukan pula perhitungan manual berdasarkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) dari confusion matrix. Perbandingan hasil perhitungan manual dengan hasil evaluasi Orange Data Mining ditunjukkan pada Tabel 2.

Tabel 2. Perbandingan Perhitungan Manual dengan Orange Data Mining

Algoritma	Accuracy	Precision	Recall / F1-Score
Decision Tree (Manual)	0,980	1,000	0,957 / 0,978
Decision Tree (Orange)	0,980	0,981	0,980 / 0,980
Random Forest (Manual)	0,984	0,978	0,991 / 0,984
Random Forest (Orange)	0,985	0,985	0,985 / 0,985

Perbedaan kecil antara hasil perhitungan manual dan Orange Data Mining terjadi karena Orange menghitung metrik menggunakan rata-rata antar kelas (macro/weighted average), sedangkan perhitungan manual menggunakan pendekatan klasifikasi biner pada satu kelas positif. Secara keseluruhan, kedua pendekatan menunjukkan bahwa Random Forest secara konsisten memperoleh nilai evaluasi yang lebih tinggi dibandingkan Decision Tree, sehingga memperkuat validitas hasil evaluasi otomatis dari Orange Data Mining.

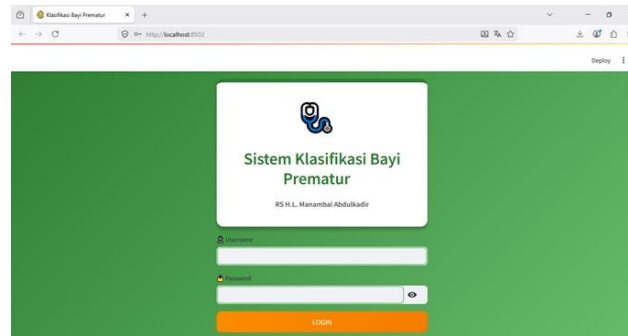
D. Pembahasan

Hasil penelitian menunjukkan bahwa algoritma Random Forest memiliki performa yang lebih baik dibandingkan Decision Tree pada seluruh metrik evaluasi. Hal ini disebabkan oleh karakteristik Random Forest yang membangun banyak pohon keputusan secara acak (ensemble) dan menggabungkan hasil prediksinya, sehingga mampu mengurangi risiko overfitting dan meningkatkan kestabilan model dibandingkan Decision Tree yang hanya mengandalkan satu struktur pohon tunggal. Temuan ini sejalan dengan penelitian [5], [6], [9], [10], [11] yang juga menunjukkan keunggulan Random Forest dibandingkan Decision Tree pada berbagai domain klasifikasi, meskipun penelitian [8] menemukan bahwa Decision Tree sedikit lebih unggul dalam nilai F1-score pada kasus tertentu meski Random Forest tetap lebih stabil. Perbedaan nilai evaluasi pada penelitian ini relatif kecil, namun tetap konsisten menunjukkan bahwa Random Forest lebih andal

digunakan dalam mengklasifikasikan kondisi kesehatan bayi prematur berdasarkan data klinis di Rumah Sakit H.L. Manambai Abdulkadir.

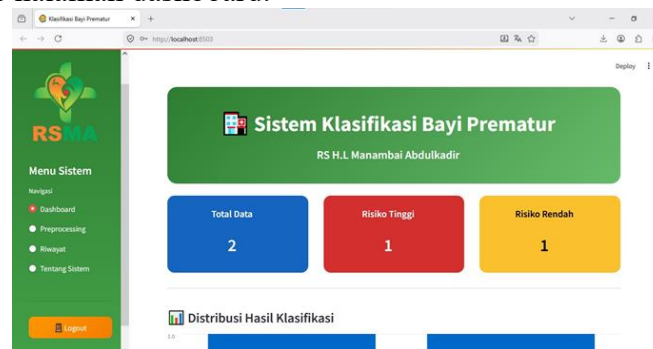
E. Implementasi Sistem

Berdasarkan hasil evaluasi, model Random Forest dipilih dan diimplementasikan ke dalam sistem klasifikasi kondisi kesehatan bayi prematur berbasis website menggunakan Python versi 3.13.3 dan framework Streamlit versi 1.45.1. Sistem menyediakan lima fitur utama, yaitu halaman login, dashboard, klasifikasi, riwayat klasifikasi, dan halaman tentang sistem.



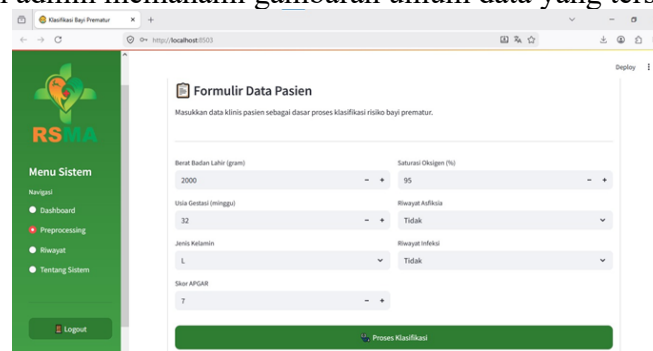
Gambar 5. Tampilan Halaman Login

Halaman login digunakan untuk melakukan autentikasi pengguna sebelum mengakses sistem. Pengguna wajib memasukkan username dan password yang valid agar dapat diarahkan ke halaman dashboard.



Gambar 6. Tampilan Halaman Dashboard

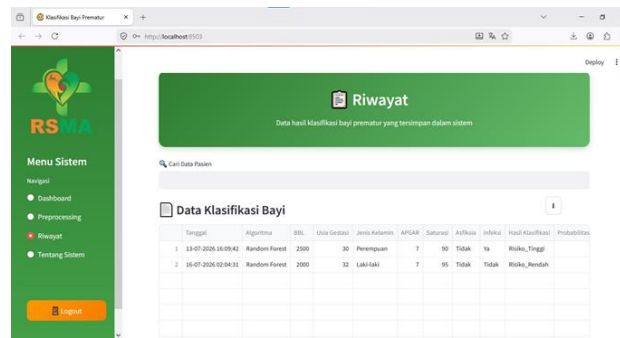
Halaman dashboard menampilkan ringkasan informasi sistem, meliputi total data, jumlah kelas Risiko_Tinggi, dan Risiko_Rendah, serta grafik distribusi hasil klasifikasi yang memudahkan admin memahami gambaran umum data yang tersimpan.



Gambar 7. Tampilan Halaman Klasifikasi

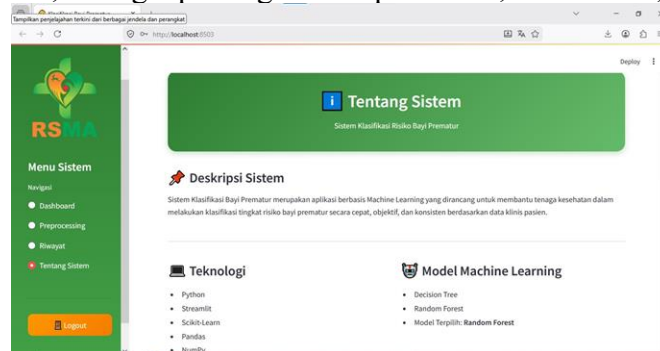
Halaman klasifikasi menyediakan form input data klinis bayi prematur, meliputi berat badan lahir, usia gestasi, jenis kelamin, skor APGAR, saturasi oksigen, kondisi asfiksia, dan kondisi infeksi. Setelah data diinput, sistem memproses data tersebut menggunakan model Random Forest dan menampilkan hasil klasifikasi beserta tingkat

kepercayaan prediksi.



Gambar 8. Tampilan Halaman Riwayat Klasifikasi

Halaman riwayat klasifikasi menampilkan seluruh data hasil klasifikasi yang tersimpan pada sistem, dilengkapi dengan fitur pencarian, unduh data, dan hapus data.



Gambar 9. Tampilan Halaman Tentang Sistem

Halaman tentang sistem menampilkan informasi mengenai tujuan pengembangan sistem serta teknologi yang digunakan, yaitu Python, Streamlit, scikit-learn, dan model Random Forest sebagai model klasifikasi utama.

F. Pengujian Sistem

Pengujian sistem dilakukan menggunakan metode Black Box Testing [15], yaitu metode pengujian yang berfokus pada kesesuaian keluaran (output) sistem terhadap masukan (input) yang diberikan tanpa memperhatikan struktur kode program. Pengujian dilakukan terhadap sepuluh skenario yang mencakup fitur login, input data klasifikasi, validasi rentang nilai input, riwayat klasifikasi, unduh dan hapus data, halaman evaluasi algoritma, hingga logout. Ringkasan sebagian skenario pengujian ditunjukkan pada Tabel 4.

Tabel 4. Ringkasan Black Box Testing

No	Skenario Pengujian	Hasil yang Diharapkan	Status
1	Login dengan username dan password benar	Sistem menampilkan halaman dashboard	Valid
2	Login dengan username/password salah	Sistem menolak akses dan menampilkan pesan kesalahan	Valid
3	Mengisi seluruh atribut data bayi dengan data valid	Sistem menampilkan hasil klasifikasi beserta tingkat kepercayaan	Valid
4	Menginput nilai di luar rentang normal (mis. BBL < 500)	Sistem menolak input dan menampilkan peringatan	Valid
5	Mengunduh dan menghapus data riwayat klasifikasi	Sistem berhasil mengunduh/menghapus data yang dipilih	Valid
6	Menekan tombol logout	Sistem mengakhiri sesi dan kembali ke halaman login	Valid

Berdasarkan hasil pengujian, seluruh sepuluh skenario Black Box Testing memperoleh status valid. Hal ini menunjukkan bahwa setiap fungsi pada sistem klasifikasi

kondisi kesehatan bayi prematur berbasis website telah berjalan sesuai dengan rancangan dan kebutuhan yang telah ditetapkan, sehingga sistem dinyatakan layak digunakan sebagai alat bantu klasifikasi di Rumah Sakit H.L. Manambai Abdulkadir.

KESIMPULAN

Berdasarkan hasil penelitian mengenai penerapan dan perbandingan algoritma Decision Tree dan Random Forest untuk klasifikasi kondisi kesehatan bayi prematur menggunakan Orange Data Mining, dapat disimpulkan bahwa algoritma Random Forest menunjukkan performa yang lebih baik dibandingkan Decision Tree. Hasil evaluasi menggunakan 10-fold cross validation menunjukkan Decision Tree memperoleh nilai accuracy 0,980, precision 0,981, recall 0,980, dan F1-score 0,980, sedangkan Random Forest memperoleh nilai accuracy, precision, recall, dan F1-score masing-masing sebesar 0,985. Berdasarkan nilai metrik yang lebih tinggi tersebut, Random Forest dipilih sebagai model terbaik untuk diimplementasikan pada sistem klasifikasi kondisi bayi prematur.

Model Random Forest berhasil diimplementasikan ke dalam sistem klasifikasi kondisi kesehatan bayi prematur berbasis website menggunakan bahasa pemrograman Python dan framework Streamlit, dengan fitur login, dashboard, klasifikasi, riwayat klasifikasi, dan halaman tentang sistem. Hasil pengujian sistem menggunakan metode Black Box Testing menunjukkan bahwa seluruh skenario pengujian pada setiap fitur memperoleh status valid, sehingga sistem dinyatakan layak digunakan sebagai alat bantu klasifikasi kondisi kesehatan bayi prematur di Rumah Sakit H.L. Manambai Abdulkadir secara lebih cepat, objektif, dan terkomputerisasi.

Penelitian ini masih memiliki keterbatasan, salah satunya jumlah dataset yang digunakan masih relatif terbatas, yaitu 200 data bayi prematur dari satu rumah sakit, sehingga performa model perlu diuji lebih lanjut pada dataset yang lebih besar dan lebih beragam. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dari beberapa rumah sakit, menambahkan atribut klinis lain seperti riwayat kesehatan ibu selama kehamilan, serta mengembangkan sistem dengan fitur multi-user, ekspor laporan dalam format PDF, dan notifikasi otomatis untuk kategori Risiko_Tinggi guna meningkatkan manfaat sistem bagi tenaga kesehatan.

DAFTAR PUSTAKA

- "Low birthweight estimates," 2015.
- "Perbandingan Tingkat Akurasi Algoritma Decision Tree dan Random Forest dalam Mengklasifikasi Penerima Bantuan Sosial BPNT di Desa Slangit," vol. 8, no. 1, pp. 127–132, 2024.
- A. A. Soleh and G. A. Rahman, "Analisis Komparasi Kestabilan Algoritma Decision Tree dan Random Forest untuk Klasifikasi Gangguan Tidur," vol. 3, no. 2, pp. 93–98, 2025.
- A. F. Azmi and A. Voutama, "Prediksi Churn Nasabah Bank Menggunakan Klasifikasi Random Forest dan Decision Tree dengan Evaluasi Confusion Matrix," vol. 13, no. 1, 2024.
- A. R. Ramadhan, N. Saefulloh, N. Utami, M. Diana, A. Aji, and P. Utomo, "Comparison of Machine Learning Models for Heart Disease Classification with Web-Based Implementation," vol. 8, no. 2, pp. 598–604, 2024.
- D. P. Sinambela, H. Naparin, M. Zulfadhilah, and N. Hidayah, "Implementasi Algoritma Decision Tree dan Random Forest dalam Prediksi Perdarahan Pascasalin," vol. 5, no. 3, pp. 58–64, 2023, doi: 10.60083/jidt.v5i3.393.
- F. Y. Pamuji and V. P. Ramadhan, "Komparasi Algoritma Random Forest dan Decision Tree untuk Memprediksi Keberhasilan Immunotherapy," vol. 7, no. 1, pp. 46–50, 2021.
- I. Akbar, F. Supriadi, and D. I. Junaedi, "Pemanfaatan machine learning di bidang kesehatan," vol. 9, no. 1, pp. 1744–1749, 2025.
- J. Demšar et al., "Orange: Data Mining Toolbox in Python," vol. 14, pp. 2349–2353, 2013.

- J. Jayadi, V. Hafizh, C. Putra, A. R. Raharja, and M. Al-husaini, "Deteksi Dini Kesehatan Mental Mahasiswa dengan Machine Learning: Perbandingan Algoritma Decision Tree dan Random Forest," vol. 16, no. 1, pp. 134–141, 2026.
- K. K. RI, "Buku Saku Hasil Survei Status Gizi Indonesia (SSGI) 2022," 2022.
- M. N. Huda, M. Burhan, A. Satibi, and H. A. Pradita, "Implementasi Black Box Testing pada Aplikasi Sistem Kasir dengan Menggunakan Teknik Equivalence Partitions," vol. 5, no. 2, pp. 120–127, 2022, doi: 10.32493/jtsi.v5i2.17645.
- M. S. Brown, "CRISP-DM: Pengantar," 2014.
- N. I. Yaman, A. R. Juwita, S. Arum, P. Lestari, and S. Faisal, "Perbandingan Kinerja Algoritma Decision Tree dan Random Forest untuk Klasifikasi Nutrisi pada Makanan Cepat Saji," pp. 184–195, 2024, doi: 10.33364/algoritma/v.21-2.1649.
- P. L. Rosida, M. Nurmalasari, and D. Krismawati, "Implementasi Decision Tree untuk Prediksi Kelahiran Bayi Prematur," vol. 14, pp. 178–186, 2024.