

ANALISIS FAKTOR PENENTU DAN PREDIKSI JENIS KONTRASEPSI PADA AKSEPTOR KB MENGGUNAKAN ALGORITMA RANDOM FOREST

Mario Edmon Gasa^{1*}, Tri Ana Setyarini², Dewi Anggraini³

^{1,2,3} Program Studi Teknik Informatika, STIKOM Uyelindo Kupang
^{1*}marioedmongasa@gmail.com, ²setyarinitriana@gmail.com, ³thewifoeh@gmail.com
 DOI: <https://doi.org/10.46880/mtk.v12i2.5960>

ABSTRACT

The Family Planning (KB) program aims to control population growth, yet the high discontinuation rate due to mismatched contraceptive choices remains a major challenge in the field. Therefore, this study aims to develop an objective contraceptive prediction model using the Random Forest algorithm to minimize the risk of program failure. The methodology involved processing 4,500 acceptor records balanced into 9 contraceptive classes with 12 demographic variables, optimized via GridSearchCV, and evaluated using 5-Fold Cross Validation. The results indicate that the model operates stably with an average accuracy of 78.87%, achieving the best performance in the Fold-1 test at 81.67%. The model also demonstrated optimal recognition for the MOP and MAL classes (F1-Score 0.98), proving the algorithm's reliability in identifying classes with highly distinctive characteristics despite data overlap challenges within the Injectable and Pill classes. Feature Importance analysis reveals that Age (22.60%), Gender (14.39%), and Age at Marriage (12.44%) are the most dominant determining factors. This prediction model is implemented in a Flask application, serving as a practical decision-support tool for healthcare workers to provide instant, transparent, and targeted contraceptive recommendations.

Keywords: Contraception, Cross Validation, Feature Importance, Prediction System, Random Forest

I. PENDAHULUAN

Keluarga Berencana (KB) merupakan program pemerintah yang berfungsi untuk menekan laju pertumbuhan penduduk dengan cara membantu individu atau pasangan suami istri dalam menghindari kelahiran yang tidak diinginkan, mengatur interval antar kelahiran, mengontrol waktu pada saat kelahiran dalam hubungan suami istri dan bisa menentukan jumlah anak dalam keluarga [1]. Untuk mencapai tujuan tersebut, program ini menyediakan beragam pilihan metode kontrasepsi yang dikelompokkan menjadi metode jangka pendek, jangka panjang, hingga permanen. Keberagaman metode ini memungkinkan akseptor untuk memilih alat kontrasepsi yang paling sesuai dengan kondisi kesehatan, rencana keluarga, dan kenyamanan masing-masing [2]. Dengan adanya program KB ini diharapkan bisa menghasilkan keluarga yang berkualitas sehingga akan mengatasi masalah kependudukan[3].

Persoalan krusial yang terus membayangi keberhasilan program KB adalah tingginya angka putus pakai (drop out) alat kontrasepsi. Fenomena ini seringkali berakar dari ketidaksesuaian antara metode yang dipilih dengan kondisi pengguna, yang dipicu oleh minimnya pengetahuan sehingga menimbulkan rasa tidak nyaman akibat efek samping. Publikasi terbaru BKKBN (2025) mencatat tingkat putus pakai nasional mencapai 20,2% pada tahun 2024. Urgensi penelitian ini semakin mendesak karena di Provinsi Nusa Tenggara Timur (NTT) sebagai lokasi studi, angkanya justru melonjak melampaui rata-rata nasional menjadi 22,6%. Temuan dari studi kasus mengonfirmasi gawatnya masalah ini, di mana terungkap fakta bahwa mayoritas akseptor KB suntik (69,1%) memilih untuk drop out [4]. Lebih lanjut, studi tersebut membuktikan bahwa faktor demografis seperti umur dan jumlah anak (paritas) menjadi penentu signifikan dalam keputusan berhenti pakai. Tentu saja, kegagalan dalam melanjutkan

kontrasepsi ini membawa konsekuensi serius, mulai dari risiko kehamilan yang tidak diinginkan hingga dampak berantai pada laju pertumbuhan penduduk yang akan memengaruhi kualitas kesejahteraan, pendidikan, dan kesehatan masyarakat.

Selama ini, konseling konvensional oleh tenaga kesehatan sering terkendala oleh terbatasnya waktu dan kompleksitas variabel profil akseptor yang harus dianalisis secara manual. Akibatnya, rekomendasi kontrasepsi menjadi kurang optimal dan berisiko memicu efek samping pada akseptor. Untuk menjembatani *gap* tersebut, diperlukan instrumen pendukung keputusan klinis berbasis komputasi yang mampu menganalisis data secara cepat dan komprehensif melalui pemanfaatan teknologi *machine learning*. Salah satu metode yang terbukti efektif dalam berbagai bidang, termasuk kesehatan, adalah algoritma Random Forest [5]. Algoritma ini termasuk dalam kategori ensemble learning yang bekerja dengan membangun banyak decision tree (pohon keputusan) dan menggabungkan hasilnya melalui proses voting untuk menghasilkan prediksi yang lebih stabil serta akurat [6]. Kelebihan utama Random Forest terletak pada kemampuannya mengatasi masalah overfitting, menangani data berskala besar dan non-linear, serta mengukur tingkat kepentingan (feature importance) dari setiap variabel, sehingga dapat diidentifikasi faktor-faktor mana yang paling berpengaruh dalam penentuan jenis kontrasepsi[7]. Dengan demikian, penerapan algoritma ini diharapkan dapat menghasilkan model prediksi yang akurat untuk merekomendasikan jenis kontrasepsi yang paling sesuai bagi calon akseptor berdasarkan karakteristik demografis, kesehatan, dan sosial-ekonomi mereka.

Penerapan algoritma Random Forest telah terbukti efektif dan akurat dalam berbagai penelitian di bidang kesehatan. Sebuah studi yang berfokus pada prediksi kelahiran bayi menunjukkan bahwa model yang dibangun

menggunakan Algoritma Random Forest, mampu mencapai akurasi sebesar 92,5% [8]. Kinerja solid dari Random Forest juga terlihat dalam penelitian lain yang membandingkan beberapa algoritma untuk memprediksi stunting pada balita. Dalam studi tersebut, Random Forest menunjukkan performa yang sangat kompetitif dengan akurasi 91% dan nilai AUC (Area Under Curve) sebesar 98% [9]. Salah satu keunggulan utama yang disorot dalam penelitian ini adalah kemampuan Random Forest untuk menyajikan analisis *feature importance*, yang berfungsi untuk mengidentifikasi faktor-faktor paling signifikan yang memengaruhi hasil prediksi.

Keandalan algoritma *Random Forest* dalam memecahkan masalah klasifikasi di bidang kesehatan telah dibuktikan secara empiris melalui beberapa studi terdahulu. Dalam kasus deteksi penyakit diabetes, algoritma ini terbukti sangat kuat dalam mengenali pola kompleks antarvariabel klinis dengan tingkat akurasi mencapai 96,88%, mengungguli metode *Neural Network* [10]. Selain itu, efisiensi dan stabilitas *Random Forest* juga teruji dalam mengklasifikasikan data pasien untuk memprediksi risiko penyakit jantung, di mana algoritma ini berhasil mencapai akurasi hingga 99%, jauh melampaui performa algoritma *K-Nearest Neighbors* (KNN) [11]. Studi komparatif lain yang mengklasifikasikan penyakit paru-paru berdasarkan data demografi pasien juga memperkuat hal ini. Meskipun akurasi *Random Forest* (91% dengan nilai AUC 0,97) berada sedikit di bawah *XGBoost* (94%), *Random Forest* terbukti unggul karena lebih cepat dalam pemrosesan, lebih mudah diinterpretasikan, serta sangat andal dalam menangani klasifikasi data medis yang kompleks [12].

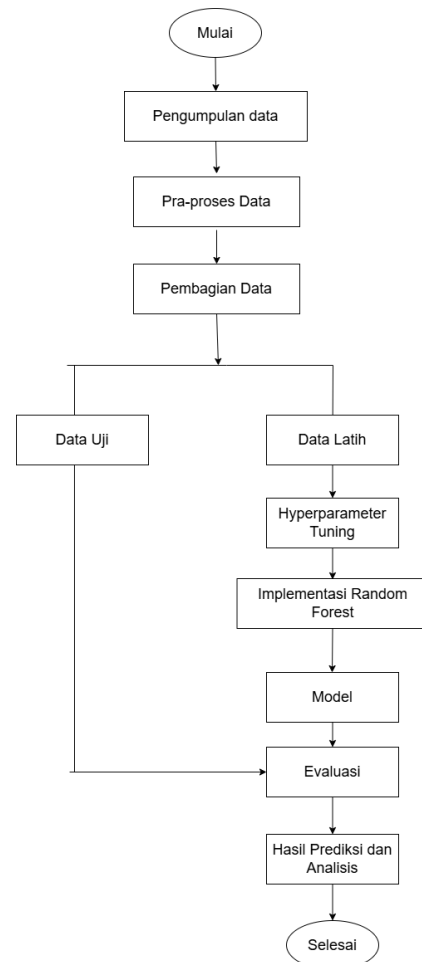
Algoritma ini merupakan bagian dari *ensemble learning* yang bekerja dengan membangun sejumlah besar pohon keputusan (*decision trees*) secara acak menggunakan teknik bootstrap aggregating atau bagging. Melalui perhitungan matematis *Gini Impurity* pada setiap percabangan dan mekanisme penggabungan keputusan akhir berbasis suara terbanyak (*majority voting*), *Random Forest* mampu mengurangi risiko *overfitting*, menangani data kompleks secara stabil, serta mengekstraksi tingkat kepentingan fitur (*feature importance*) untuk menganalisis faktor penentu secara transparan [16].

Rekam jejak performa yang solid dalam menganalisis data medis tersebut menjadi landasan kuat bagi penerapan algoritma Random Forest pada penelitian ini. Dengan menggunakan data resmi dari BKKBN yang telah memenuhi standar etika privasi, penelitian ini berkontribusi menganalisis faktor penentu dan memprediksi pilihan 9 (sembilan) kelas jenis kontrasepsi, serta mengintegrasikan model tersebut ke dalam aplikasi berbasis web menggunakan *Flask* sebagai alat bantu praktis bagi tenaga kesehatan.

II. METODE

Prosedur Penelitian

Tahapan penelitian ini disusun secara sistematis, mulai dari pengumpulan data hingga implementasi algoritma *Random Forest*. Untuk menjamin keterukuran alur kerja agar dapat direproduksi, keseluruhan proses tersebut divisualisasikan pada Gambar 1 berikut:



Gambar 1 Flowhart Penelitian

1. Pengumpulan dan Pra-proses Data

Penelitian ini diawali dengan mengumpulkan data sekunder dari BKKBN Provinsi NTT yang berisi rekam jejak akseptor KB. Data mentah tersebut selanjutnya melalui tahapan pra-proses terukur yang meliputi pembersihan data (*data cleaning*) dari duplikasi dan nilai kosong, transformasi variabel kategorikal menjadi numerik (*label encoding*), serta penyeimbangan data (*data balancing*). Proses penyeimbangan ini mendistribusikan data secara merata ke dalam 9 kelas jenis kontrasepsi dengan porsi masing-masing 500 sampel, sehingga menghasilkan total himpunan 4.500 data bersih yang sepenuhnya kompatibel untuk perhitungan algoritma.

2. Pembagian Data dan Hyperparameter Tuning

Pemilihan 9 kelas target pada penelitian ini didasarkan pada pengelompokan resmi metode kontrasepsi BKKBN. Guna mencegah model condong (*bias*) pada kelas mayoritas, dilakukan mekanisme penyeimbangan data menggunakan teknik SMOTE sehingga setiap kelas memiliki porsi identik sebesar 500 sampel. Dataset yang telah seimbang ini kemudian didistribusikan menjadi 80% data latih dan 20% data uji secara acak terstruktur dengan menetapkan parameter *seed* (*random_state= 42*) untuk memastikan reproduksibilitas pembagian data. seluruh tahapan komputasi ini, mulai dari pemrosesan data hingga pemodelan, diimplementasikan menggunakan pustaka mesin pembelajaran *Scikit-Learn* dalam bahasa pemrograman *Python* [17].

Selanjutnya, model dioptimasi menggunakan *hyperparameter tuning* dengan *GridSearchCV* yang divalidasi

secara internal melalui *5-Fold Cross Validation*. Prosedur validasi silang ini membagi data latih menjadi lima lipatan (*fold*) independen guna memastikan model tidak menghafal data (*overfitting*) serta menghindari bias evaluasi terhadap subset data tertentu.

3. Implementasi Random Forest dan Evaluasi Model

Setelah parameter terbaik dikonfigurasi, algoritma *Random Forest* diimplementasikan untuk mempelajari pola data latih. Pada tahap prediksi, ambang keputusan (*decision threshold*) model didasarkan pada mekanisme suara terbanyak (*majority voting*), di mana probabilitas ambang batas klasifikasi berada pada nilai $>0,5$ dari seluruh pohon keputusan (*decision trees*) yang terbentuk. Performa model akhir kemudian diuji menggunakan 20% data uji melalui *Confusion Matrix* untuk menghitung metrik kinerja utama, yaitu: Akurasi (persentase ketepatan prediksi model secara keseluruhan), Presisi (rasio ketepatan prediksi terhadap data aktual untuk menghindari *false positive*), *Recall* (kemampuan model dalam mengenali/menemukan seluruh data pada setiap kelas), dan *F1-Score* (rata-rata harmonis antara presisi dan *recall* untuk mengatasi ketidakseimbangan kelas). Klaim keandalan kinerja sistem klasifikasi diukur berdasarkan perolehan nilai akhir metrik-metrik tersebut, di mana nilai yang mendekati 1 (100%) merepresentasikan kemampuan prediksi yang optimal.

4. Hasil Prediksi dan Analisis Faktor Penentu

Tahap akhir dari alur ini memproduksi dua luaran fungsional utama dari algoritma *Random Forest*, yakni prediksi jenis kontrasepsi yang paling direkomendasikan berdasarkan mekanisme suara terbanyak (*majority voting*). Selain itu, sistem mengekstraksi metrik *Feature Importance* untuk menganalisis variabel demografis yang paling dominan. Ekstraksi ini didasarkan pada seberapa besar suatu variabel mampu menurunkan tingkat ketidakmurnian data pada saat pembentukan pohon keputusan. Adapun formula matematis yang digunakan oleh algoritma *Random Forest* untuk menghitung nilai *Gini Impurity* pada setiap percabangan simpul adalah sebagai berikut:

$$Gini_t = 1 - \sum_{i=1}^n p_i^2 \quad (1)$$

dimana n merupakan jumlah kelas target klasifikasi (jenis kontrasepsi) dan p_i adalah proporsi sampel yang termasuk dalam kelas ke- i pada simpul tersebut. Variabel yang secara matematis menghasilkan penurunan nilai *Gini* paling besar akan mendapatkan skor *Feature Importance* tertinggi, sehingga sistem dapat mengidentifikasi faktor penentu keputusan akseptor secara transparan dan objektif.

III. HASIL DAN PEMBAHASAN

Implementasi Sistem

Sistem prediksi jenis kontrasepsi dibangun menggunakan algoritma *Random Forest* yang dioptimasi untuk mengklasifikasikan 12 variabel demografi akseptor. Model dilatih menggunakan 4.500 rekam jejak data historis dari BKKBN yang telah melalui proses penyeimbangan (data balancing) ke dalam 9 kelas kontrasepsi, di mana setiap kelas memiliki porsi merata sebanyak 500 data. Sebelum memasuki tahap pelatihan utama, model melewati proses *hyperparameter tuning* menggunakan

GridSearchCV yang menetapkan konfigurasi performa paling stabil pada parameter $n_estimators$ sebanyak 300, max_depth di level 25, dan $min_samples_split$ pada angka 5.

Evaluasi keandalan model divalidasi secara menyeluruh menggunakan *5-Fold Cross Validation*. Model cerdas ini kemudian diintegrasikan ke dalam antarmuka aplikasi berbasis web menggunakan *framework Flask* agar dapat memberikan rekomendasi prediksi dan visualisasi faktor penentu secara instan kepada pengguna.

Pengujian Sistem

Pengujian model bertujuan untuk mengukur performa prediksi menggunakan variasi skenario pembagian data sekaligus mengevaluasi variabel yang paling dominan.

1. Hasil Evaluasi K-Fold Cross-Validation

Algoritma *Random Forest* terbukti andal dan stabil dengan pencapaian rata-rata akurasi sebesar 78,87% melintasi lima skenario pengujian. Rincian akurasi pada setiap lipatan (*fold*) menunjukkan bahwa *Fold-2* mencatatkan akurasi 77,33%, *Fold-3* sebesar 77,44%, *Fold-4* sebesar 79,56%, dan *Fold-5* sebesar 78,33%. Performa puncak atau titik optimal klasifikasi dicapai pada pengujian *Fold-1* dengan tingkat akurasi tertinggi mencapai 81,67%.

Tabel 1. Hasil Pengujian dengan K-Fold Cross Validation

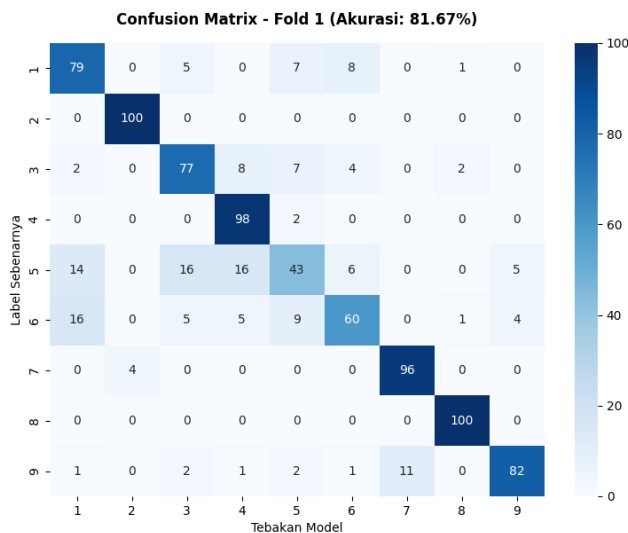
No	K-Fold	Akurasi
1	<i>Fold-1</i>	81,67%
2	<i>Fold-2</i>	77,33%
3	<i>Fold-3</i>	77,44%
4	<i>Fold-4</i>	79,56%
5	<i>Fold-5</i>	78,33%

2. Analisis Evaluasi Confusion Matrix

Evaluasi *Confusion Matrix* pada model terbaik (*Fold-1*) yang menghasilkan akurasi tertinggi sebesar 81,67%, tingkat keberhasilan prediksi dapat dilihat pada kotak-kotak diagonal utama (*True Positive*). Model menunjukkan performa pengenalan yang sangat presisi pada kelas MOP (2) dan MAL (8) yang masing-masing berhasil ditebak sempurna sebanyak 100 data, disusul kelas Implant (4) sebanyak 98 data.

Meskipun demikian, model masih mengalami kebingungan (misklasifikasi) yang cukup signifikan pada kontrasepsi kelas Suntik (5) dan Pil (6). Sebagai contoh, terdapat 16 data aktual kelas Suntik yang keliru diprediksi sebagai IUD (3) dan 16 data diprediksi sebagai Implant (4). Selain itu, 16 data aktual kelas Pil juga salah diklasifikasikan menjadi kelas MOW (1).

Tingkat kesalahan pada sel di luar diagonal ini mengindikasikan bahwa batasan antar kelas pada metode tersebut masih kurang tegas akibat tingginya kemiripan profil demografi (*data overlap*) antar akseptornya.



Gambar 2. Confusion Matrix

3. Analisis Laporan Klasifikasi

Evaluasi model *Random Forest* melalui *5-fold cross-validation* mencatatkan rata-rata akurasi sebesar 78,87% dalam mengklasifikasikan 9 target metode kontrasepsi, dengan rincian metrik Presisi, *Recall*, dan *F1-Score* per kelas disajikan pada Tabel 1. Berdasarkan model terbaik (*Fold-1*), algoritma mencapai tingkat pengenalan paling optimal pada metode MOP dan MAL dengan nilai *F1-Score* menyentuh 0,98. Sebaliknya, performa terendah tercatat pada kelas Suntik (*F1-Score* 0,51) dan Pil (*F1-Score* 0,67).

Penurunan metrik pada kedua kelas tersebut mempertegas bahwa terjadinya *overlap* prediksi disebabkan oleh kemiripan distribusi profil demografi antara akseptor Suntik dan Pil, sehingga algoritma kesulitan dalam memberikan batasan klasifikasi yang mutlak. Kasus *overlap* atau *false positive* semacam ini berisiko secara klinis jika sistem salah merekomendasikan kontrasepsi, sehingga aplikasi *web* yang dibangun harus diposisikan secara ketat sebagai Sistem Pendukung Keputusan yang tetap membutuhkan validasi akhir tenaga medis.

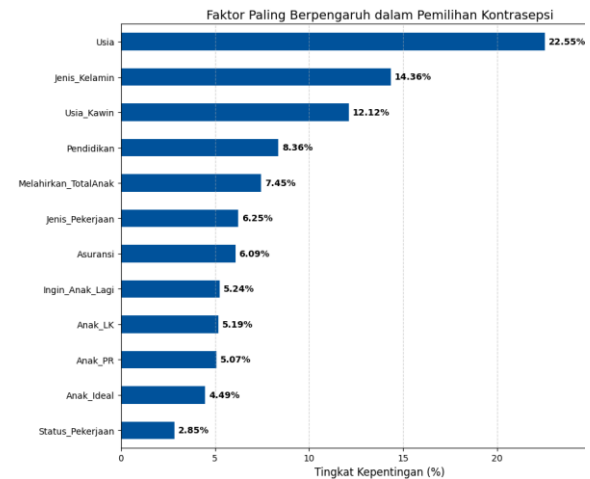
Lebih lanjut, kemampuan generalisasi model ini masih dibatasi oleh ketiadaan variabel rekam medis klinis dalam dataset sekunder BKKBN, sehingga prediksi model saat ini murni bertumpu pada probabilitas tren sosio-demografi dan belum sepenuhnya mampu memfilter kontraindikasi medis spesifik dari akseptor.

Tabel 2. Laporan Klasifikasi

Kelas_Kontrasepsi	Precision	Recall	F1-Score	(Support)
1 (MOW)	71 %	79 %	74 %	100
2 (MOP)	96 %	100 %	98 %	100
3 (IUD)	73 %	77 %	75 %	100
4 (Implant)	77 %	98 %	86 %	100
5 (Suntik)	61 %	43 %	51 %	100
6 (Pil)	76 %	60 %	67 %	100
7 (Kondom)	90 %	96 %	93 %	100
8 (MAL)	96 %	100 %	98 %	100
9 (Tradisional)	90 %	82 %	86 %	100

4. Analisis Faktor Penentu (Feature Importance)

Untuk melacak variabel yang memandu pengambilan keputusan sistem, dilakukan ekstraksi tingkat kepentingan fitur (*Feature Importance*). Ekstraksi ini mengonfirmasi bahwa Usia (22,60%), Jenis Kelamin (14,39%), dan Usia Kawin (12,44%) merupakan tiga faktor penentu paling dominan. Penemuan ini secara ilmiah menegaskan bahwa pertimbangan medis dan biologis Fedasar menjadi landasan utama yang memengaruhi akseptor dalam memilih jenis kontrasepsi.



Gambar 3. Feature Importance

Kebijakan dan Kekurangan Sistem

Sistem prediksi berbasis *Random Forest* ini memiliki keunggulan pada kecepatan proses analisis instan dan objektivitas rekomendasi yang murni berbasis pola data historis, sehingga sangat efektif dijadikan alat bantu bagi akseptor baru. Tingkat transparansi sistem juga sangat baik karena dilengkapi dengan visualisasi grafik *feature importance* yang memberikan penjelasan transparan mengenai variabel apa yang paling memengaruhi hasil prediksi.

Secara keilmuan, hasil riset ini memberikan kontribusi akademis berupa pemanfaatan teknologi *machine learning* untuk mendukung efektivitas program kesehatan masyarakat, khususnya dengan membuktikan kemampuan algoritma klasifikasi dalam memetakan dan mengukur faktor-faktor demografis penentu keputusan akseptor secara kuantitatif. Meskipun demikian, sistem ini memiliki batasan pada cakupan analisis yang hanya bekerja berdasarkan profil demografi dan sosial akseptor. Karena variabel medis yang mendalam (seperti hasil laboratorium atau riwayat penyakit kronis) tidak dimasukkan, sistem memotret kebutuhan akseptor murni dari sisi latar belakang kehidupannya saja.

IV. KESIMPULAN

Penelitian ini menyimpulkan bahwa model *Random Forest* berhasil memprediksi 9 kelas metode kontrasepsi dengan kinerja stabil, mencapai rata-rata akurasi 78,87% (tertinggi 81,67% pada *Fold-1*), meskipun masih terdapat *overlap* pada metode jangka pendek (Suntik dan Pil) akibat kemiripan profil akseptor. Analisis *Feature Importance* mengungkap bahwa pola pilihan akseptor secara dominan dipengaruhi oleh faktor demografi, dengan variabel penyumbang bobot terbesar adalah Usia (22,60%), Jenis Kelamin (14,39%), dan Usia Kawin

(12,44%). Sebagai rekomendasi praktis, aplikasi *web* pendukung keputusan yang diusulkan ini memerlukan validasi klinis langsung oleh tenaga kesehatan di lapangan sebelum dieksekusi untuk deployment massal.

Melalui pendekatan ini, penelitian diharapkan memberikan kontribusi ilmiah dalam pengembangan algoritma prediktif multikelas di bidang medis, sekaligus menghadirkan manfaat sosial berupa optimalisasi layanan konseling Keluarga Berencana yang lebih efisien dan tepat sasaran.

V. REFERENSI

- [1] Badan Pusat Statistik Provinsi Nusa Tenggara Timur, "Statistik Sosial dan Kependudukan Provinsi Nusa Tenggara Timur *Social and Demography Statistics of Nusa Tenggara Timur Province*," 2022. Accessed: Oct. 08, 2025. [Online]. Available: <https://ntt.bps.go.id/id/publication/2023/06/15/9318df0e993ae76e1edcae45/statistik-sosial-dan-kependudukan-provinsi-nusa-tenggara-timur-2022.html>
- [2] V. A. Simanjuntak and R. Hasibuan, "Faktor Penggunaan Metode Kontrasepsi Jangka Panjang (MKJP)," *Jurnal Kebidanan Malakbi*, vol. 5, no. 2, p. 66, Aug. 2024, doi: 10.33490/b.v5i2.786.
- [3] K. E. Haseli, A. A. Adu, and D. S. Tirra, "Faktor-Faktor Yang Berhubungan Dengan Penggunaan Metode Kontrasepsi Jangka Panjang Pada Wanita Usia Subur," *Artikel Penelitian Jurnal Kesehatan*, vol. 12, no. 2, 2023, doi: 10.37048/kesehatan.v12i2.131.
- [4] J. A. Lubis, L. Barus, P. Studi, S.-1 Kebidanan, S. Tingi, and I. Kesehatan, "Faktor-Faktor Yang Berhubungan Dengan Kejadian *Drop Out* Alat Kontrasepsi Suntik Di Poskesdes Sion Timur II Tahun 2020," 2020. Accessed: Oct. 08, 2025. [Online]. Available: <https://www.mid-wifery.jurnalsenior.com/index.php/ms/article/view/39/40>
- [5] M. Santoso, M. Zydan K, I. Iwan, and M. Mudrika, "Prediksi Risiko Penyakit Jantung Sederhana Menggunakan Algoritma *Random Forest Classifier* Dengan Data Gaya Hidup Siswa," *Jurnal Manajemen Informatika Jayakarta*, vol. 5, no. 4, p. 311, Dec. 2025, doi: 10.52362/jmijayakarta.v5i4.2102.
- [6] A. Y. Agusyl and F. Firmansyah, "Prediksi Penyakit Jantung Menggunakan Algoritma *Random Forest*," *Jurnal Minfo Polgan*, vol. 12, no. 2, Nov. 2023, doi: 10.33395/jmp.v12i2.13214.
- [7] M. Mahendra Alvanof and R. Kesuma Dinata, "Penerapan Algoritma *Random Forest* dalam Deteksi dan Klasifikasi *Ransomware*," 2024. Accessed: Jul. 10, 2026. [Online]. Available: <https://jurnal.uniki.ac.id/index.php/jet/article/view/488>
- [8] Herianto, Ahmad Rofi'i, and Ribut Julianto, "Prediksi Kelahiran Bayi Berbasis Sistem Informasi Menggunakan *Machine Learning*," 2025, Accessed: Oct. 10, 2025. [Online]. Available: <https://jurnal.univwirabuana.ac.id/index.php/jukes/article/view/213>
- [9] N. R. Febriyanti, K. Kusriani, and A. D. Hartanto, "Analisis Perbandingan Algoritma SVM, *Random Forest* dan *Logistic Regression* untuk Prediksi Stunting Balita," *Edumatic: Jurnal Pendidikan Informatika*, vol. 9, no. 1, pp. 149–158, Apr. 2025, doi: 10.29408/edumatic.v9i1.29407.
- [10] M. Aqil Zidane *et al.*, "Penilaian Komparatif Metode Klasifikasi *Neural Network* Dan *Random Forest* Untuk *Knowledge Discovery* Pada Penyakit Diabetes," 2025. [Online]. Available: <https://www.kaggle.com/datasets/alextebo>
- [11] E. Sahelvi, P. Cikita, R. M. Sapitri, R. Rahmaddeni, and L. Efrizoni, "Perbandingan Algoritma *K-Nearest Neighbors* dan *Random Forest* untuk Rekomendasi Gaya Hidup Sehat dalam Mencegah Penyakit Jantung," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 3, pp. 830–840, Jun. 2025, doi: 10.57152/malcom.v5i3.1972.
- [12] R. Harahap, M. Irpan, M. Azzuhri Dinata, and L. Efrizoni, "Perbandingan Algoritma *Random Forest* Dan *Xgboost* Untuk Klasifikasi Penyakit Paru-Paru Berdasarkan Data Demografi Pasien," 2024. Accessed: Jul. 10, 2026. [Online]. Available: <https://ejournal.pppmitpa.or.id/index.php/betrik/article/view/231>
- [13] Kementerian Kesehatan Republik Indonesia, *Pedoman Pelayanan Kontrasepsi*. Jakarta: Kementerian Kesehatan RI, 2021. Accessed: Nov. 19, 2025. [Online]. Available: <https://repository.kemkes.go.id/book/571>
- [14] F. Afifah Nurullah, "Continuing Medical Education Akreditasi PB IDI-2 SKP Perkembangan Metode Kontrasepsi di Indonesia." doi: 10.55175/cdk.v48i3.1335.
- [15] Leo Breiman, "randomforest2001," Jan. 2001.
- [16] N. Nur, S. Situju, H. Setiawan, and T. Kristanto, *Machine Learning*. 2025. [Online]. Available: <https://www.researchgate.net/publication/394364101>
- [17] F. Pedregosa Fabianpedregosa *et al.*, "Scikit-learn: *Machine Learning* in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos Pedregosa, Varoquaux, Gramfort Et AL. Matthieu Perrot," 2011. [Online]. Available: <http://scikit-learn.sourceforge.net>.