



## Implementasi Algoritma *Decision Tree* dan *Random Forest* dalam Prediksi Perdarahan Pascasalin

Dewi Pusparani Sinambela<sup>1✉</sup>, Husni Naparin<sup>2</sup>, Muhammad Zulfadhilah<sup>3</sup>, Nurul Hidayah<sup>4</sup>

<sup>1,4</sup>Diploma Kebidanan, Universitas Sari Mulia

<sup>2,3</sup>Teknologi Informasi, Universitas Sari Mulia

[dewipusparani@unism.ac.id](mailto:dewipusparani@unism.ac.id)

### Abstrak

Perdarahan Postpartum (PPP) merupakan salah satu kegawatdaruratan pada persalinan yang dapat menyebabkan kematian di negara maju dan negara berkembang. Salah satu pencegahan terjadinya PPP dengan melakukan prediksi pada ibu bersalin dengan mempertimbangkan faktor faktor risiko menggunakan pendekatan model Machine Learning (ML). Algoritma *Random Forest* (RF) dan *Decision Tree* (DT) merupakan algoritma yang digunakan dalam prediksi kejadian PPP. Tujuan dari penelitian ini adalah mengembangkan kinerja dari Algoritma RF dan Algoritma RF untuk mengklasifikasi kejadian PPP. Hasil analisis Berdasarkan hasil analisis univariat yang ditunjukkan pada tabel 1 didapatkan ibu yang memiliki paritas > 4 sebanyak 102 orang (20,4%), jarak kehamilan ibu yang  $\leq 2$  tahun sebanyak 310 orang (62%), ibu pasca bersalin yang mengalami anemia sebanyak 124 orang (24,8%), ibu yang melahirkan bayi makrosomia sebanyak 60 orang (12 %), ibu yang mengalami komplikasi persalinan sebanyak 229 orang (45,8 %), ibu yang mengalami kehamilan ganda sebanyak 16 orang (3,2%), umur ibu yang berisiko sebanyak 132 orang (26,4%). Perbandingan tingkat akurasi algoritma RF mencapai 0,830 dibandingkan dengan algoritma DT sebesar 0.820, AUC RF 0.74. Hal ini menunjukkan bahwa Algoritma RF mempunyai performace metric lebih naik dibandingkan dengan algoritma DT. Algoritma Random Forest dapat dianggap sebagai salah satu algoritma representatif ML, yang dikenal karena kemudahannya dan efektivitasnya.

**Kata Kunci:** *Decision Tree*, Perdarahan Postpartum, Prediksi, Machine Learning, *Random Forest*.

JITD is licensed under a Creative Commons 4.0 International License.



### 1. Pendahuluan

Perdarahan Postpartum (PPP) adalah kondisi perdarahan yang terjadi setelah seorang wanita melahirkan bayinya [1], [2]. Perdarahan merupakan salah satu masalah utama dalam bidang obstetri dan dapat menjadi penyebab kematian maternal jika tidak ditangani dengan cepat dan efektif [3]. Bersama dengan preeklampsia/eklampsia (tekanan darah tinggi saat hamil) dan infeksi PPP merupakan salah satu dari tiga penyebab utama kematian maternal baik di negara maju maupun di negara sedang berkembang. Angka kematian ibu di Indonesia pada tahun 2017 adalah sebesar 305 per 100.000 kelahiran hidup. Ini menunjukkan penurunan dari tahun 2012 yang mencapai 359 per 100.000 kelahiran hidup, tetapi masih meningkat dibandingkan dengan angka tahun 2007 yang mencapai 228 per 100.000 kelahiran hidup. Meskipun ada penurunan sejak 2012, angka kematian ibu yang tinggi adalah masalah serius dalam sistem perawatan kesehatan di Indonesia [4]–[7].

Perdarahan pascapersalinan saat masuk persalinan dapat diprediksi dengan kemampuan diskriminatif yang sangat baik menggunakan model Machine Learning (ML) [8]. Model pembelajaran mesin adalah pendekatan yang kredibel untuk meningkatkan prediksi PPP dengan akurasi tinggi. Penerapan klinis lebih lanjut diperlukan, yang dapat membantu penyedia layanan kesehatan untuk bersiap dan melakukan triase pada perempuan berisiko. Model ML yang dapat diinterpretasikan memungkinkan tenaga kesehatan untuk membuat keputusan yang masuk akal dan berdasarkan data untuk memberikan keputusan yang dipersonalisasi yang pada akhirnya dapat meningkatkan kualitas layanan di bidang kesehatan [9] [10].

Untuk mengurangi risiko PPH (Perdarahan Pasca Persalinan) pada wanita saat melahirkan, dokter kandungan harus mempertimbangkan faktor-faktor risiko yang telah ada dan menggunakan skema stratifikasi risiko untuk mengestimasi kemungkinan terjadinya PPH. Dalam konteks peningkatan upaya pencegahan yang ditekankan, pedoman manajemen PPP memiliki sedikit alat yang secara akurat dapat memprediksi wanita yang memiliki risiko tertinggi mengalami PPP. Perkembangan terbaru di bidang ilmu komputer, terutama dalam kecerdasan buatan, telah mendorong kemajuan dalam bidang ini. Salah satu perbedaan yang paling mencolok adalah bahwa statistik konvensional tergantung pada model, sementara AI dan Machine Learning (ML) didasarkan pada data, tanpa memerlukan pengetahuan sebelumnya tentang hubungan antara data dan hasil. Penggunaan AI dan ML dalam pengaturan perawatan kesehatan perempuan telah berkembang pesat dalam beberapa tahun terakhir [11]–[13]. Penelitian yang berkaitan dengan faktor-faktor yang terkait dengan risiko perdarahan postpartum (PPH) telah menjadi subjek penelitian yang mendalam dengan menggunakan metode statistik konvensional [14], [15]. Model-

model seperti regresi logistik atau model regresi Lasso telah diterapkan untuk meramalkan risiko PPH. Secara spesifik, model regresi Lasso atau logistik telah menunjukkan hasil yang baik dalam hal kemampuan diskriminatif [12]. Berdasarkan karakteristik klinis ibu dan riwayat medisnya, skor risiko telah digunakan untuk memprediksi kemungkinan terjadinya PPH [13], [14]. Meskipun metode prediksi konvensional telah terbukti efektif, performa mereka dapat ditingkatkan melalui pendekatan yang lebih canggih atau teknik statistik yang lebih mutakhir.

Menurut Nasiba et al, (2021) [16] menyatakan model ML merupakan teknik yang signifikan untuk mewujudkan Kecerdasan Buatan. Algoritma Random Forest dapat dianggap sebagai salah satu algoritma representatif ML, yang dikenal karena kemudahannya dan efektivitasnya. Random Forest (RF) adalah teknik klasifikasi grup paling diterima karena memiliki fitur-fitur yang sangat baik seperti Pengukuran Pentingnya Variabel, Kesalahan Out-of-bag, Proksimitas, dll. Saat ini, ini digunakan dalam klasifikasi baru, deteksi intrusi, penyaringan informasi konten, dan analisis sentimen, itulah mengapa ada beragam aplikasi dalam pengolahan gambar. Selain itu, penelitian Umri et al, (2022) [17] [18] melakukan pengujian metode K-Nearest dan Random Forest untuk melihat akurasi dari prediksi keberhasilan data yang diujikan, diperoleh hasil yaitu akurasi dari metode KNN sebesar 90.00 % dan metode Random Forest dengan tingkat akurasi 85,50 %. Dari hasil yang diperoleh terhadap pengujian yang telah dilakukan, diketahui bahwa metode Random Forest merupakan metode yang lebih baik dibandingkan dengan K-Nearest Neighbor dalam memprediksi akurasi pada Immunotherapy Dataset. Disamping itu, menurut penelitian Raza et al, (2022) dari data kehamilan dari 600 sampel dikumpulkan dari tiga pusat medis di Bangalore. Model DT yang digunakan berhasil mencapai akurasi sebesar 71%. Nilai kesalahan relatif mutlak adalah 99%, dan nilai akar dari kesalahan kuadrat terkait juga sebesar 99% untuk klasifikasi DT memprediksi komplikasi kesehatan dengan menggunakan dua klasifikasi berbasis pembelajaran mesin [19] [20]. Venkatesh et al [21] juga menggunakan pendekatan RF sebagai metode dalam memprediksi PPP.

Berdasarkan kajian studi literature, peneliti memiliki tujuan mengembangkan kinerja dari Algoritma Random Forest dan Decision Tree. Algoritma RF dan DF digunakan untuk mengklasifikasi kejadian Perdarahan Pascasalin. Dengan alasan tersebut, algoritma yang digunakan merujuk pada *Machine Learning classifier*. Tujuan jangka panjangnya adalah memungkinkan membantu tenaga medis untuk mengidentifikasi pasien yang berisiko tinggi mengalami PPH dan membimbing pengambilan keputusan klinis.

## 2. Metode

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian kuantitatif merupakan jenis penelitian untuk mendapatkan gambaran yang akurat.

### 2.1 Data set

Data set dalam penelitian ini adalah seluruh ibu pasca bersalin yang tercatat di buku register pada ruang nifas dan ruang bersalin di RSUD. dr. H. Moch Ansari Saleh Banjarmasin sebanyak 3258 orang. Sampel penelitian ini adalah sebagian seluruh ibu yang mengalami PPP dan sebagian ibu yang tidak mengalami PPP di RSUD. dr. H. Moch Ansari Saleh Banjarmasin

### 2.2 Data Analisis

Studi ini menggunakan pendekatan machine learning dengan menggunakan algoritma DT dan RF model, kemudian membandingkan kedua algoritma dengan menggunakan cross fold validation untuk membangun kinerja system dalam memprediksi kejadian Perdarahan Postpartum. Algoritma yang digunakan meliputi:

#### 2.2.1 Random Forest

Random Forest merupakan suatu set dari decision trees yang dibangun dengan sampel yang dipilih secara acak tetapi memiliki peraturan membelah simpul yang berbeda. Model ini bekerja dengan menggunakan subset dari suatu fitur pada setiap pohon, kemudian mencoba mencari ambang batas terbaik didalam memisahkan data. Sehingga hasilnya akan memiliki banyak pohon yang dilatih dengan cara yang lebih lemah dan masing-masing akan menghasilkan prediksi yang berbeda. Hasil tersebut kemudian dapat diinterpretasi dengan dua cara, yang paling umum adalah berdasarkan suara terbanyak sehingga akan di pertimbangkan sebagai kelas yang benar. Tetapi scikit-learn melakukan implementasi algoritmanya berdasarkan rata-rata dari hasil tersebut sehingga menghasilkan prediksi yang sangat akurat. Sehingga walaupun secara teoritis berbeda, rata-rata probabilitas dari Random Forest yang terlatih tidak dapat sangat berbeda dari sebagian besar prediksi (jika tidak, harus ada titik stabil yang berbeda); oleh karena itu kedua metode tersebut sering kali mengarah ke hasil yang sebanding. Untuk model Random Forest dengan scikit-learn ada beberapa parameter yang dapat diatur, seperti untuk menentukan jumlah pohon yang ingin model ini bangun(n\_estimators) [22] [23] [24].

Ada berbagai cara untuk mengukur ketidakakuratan dan algoritma yang berbeda untuk mewujudkan pohon tersebut. Algoritme penerapan pohon yang berbeda mengikuti ukuran ketidakakuratan yang berbeda pula. Ukuran ketidakakuratan dipilih untuk mengurangi ketidakakuratan yang dicapai oleh pemisahan pada node yang diamati. Jadi kita dapat menghitung throughput gabungan s pada sebuah node sebagai :

$$\Delta \eta(s, t) = \eta(t) - K(l) \eta(l) - K(r) \eta(r) \quad (1)$$

Dimana  $K(l)$  adalah proporsi sisi kiri dari kasus-kasus yang dipecah, dan untuk sisi kanan adalah  $K(r)$ . Indeks Gini adalah salah satu teknik terpenting untuk mengukur ketidakakuratan. Untuk kasus dua kelas dapat dihitung sebagai berikut :

$$\eta(t) = v(0|t) v(1|t) = v(0|t)(1-v(0|t)) = v(1|t)(1-v(1|t)) \quad (2)$$

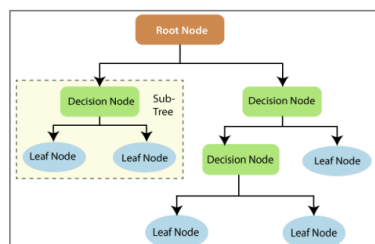
Dimana  $v(0|t)$  or  $v(1|t)$  adalah frekuensi kelas label dalam node. Generalisasi indeks Gini dengan lebih dari dua kelas dapat dihasilkan sebagai berikut :

$$\eta(t) = \sum_j v(j|t) (1 - v(j|t)) \quad (3)$$

Dimana  $v(j|t)$  adalah frekuensi rrelative pada kelas j di node t.

### 2.2.2 Decision Tree

Salah satu metode yang kerap digunakan dalam proses data mining adalah penerapan sistem pengklasifikasi. Dalam konteks penambangan data, algoritma klasifikasi memiliki kemampuan untuk mengelola volume informasi yang besar. Pendekatan ini bermanfaat untuk melakukan prediksi terkait kategori atau label kelas, menyusun pengetahuan berdasarkan data pelatihan dan labelnya, serta melakukan klasifikasi pada data yang baru diperoleh. Di dalam domain pembelajaran mesin, terdapat beragam algoritma klasifikasi yang tersedia. Dalam naskah ini, perhatian utamanya difokuskan pada konsep umum dari algoritma pohon keputusan. Ilustrasi dalam Gambar 1 memvisualisasikan struktur dari pohon keputusan (DT) [25][26].



Gambar 1 Decision Tree

Seperti yang terlihat pada Gambar 1, kotak di bagian atas pohon disebut sebagai node akar, dan di bawahnya disebut sebagai node. Di setiap kotak terdapat proses pengambilan keputusan yang menghasilkan dua cabang ke kiri dan kanan. Pengambilan keputusan dilakukan dengan menguji suatu variabel.

Secara umum dalam pembuatan DT ada beberapa algoritma yang umum digunakan, salah satunya disebut CHART (Classification and Regression Trees) yang diterapkan dalam scikit learning. Chart merupakan turunan dari algoritma lain bernama C4.5, sedangkan C4.5 sendiri merupakan turunan dari algoritma ID3 yang dibuat pada tahun 1980. Pada dasarnya algoritma ID3 membagi data menjadi dua kelompok berdasarkan atribut yang ada pada data tersebut, dengan cara mengukur suatu bilangan yang disebut entropi. Entropi merupakan rumus untuk menghitung homogenitas atribut (A) suatu sampel data (S) dengan rumus:

$$\text{Entropy (S)} = \sum_{i=1}^n - p_i * \log_2 p_i \quad (4)$$

Selanjutnya,

$$\text{Entropy (S)} = \sum_{i=1}^n - p(i|s) \log_2 p(i|s) \quad (5)$$

Kemudian Gain (S,A) merupakan informasi Gain of attribut A dalam sampel koleksi S

$$\text{Gain (S,A)} = \text{Entropy (S)} - \sum_{i=1}^n \frac{|S_i|}{|S|} \text{Entropy}(S_i) \quad (6)$$

Gain informasi dengan nilai index Gini

$$\text{Gini (t)} = 1 - \sum_{i=1}^n [p(i|s)]^2 \quad (7)$$

$$\text{Gain (S,A)} = \text{Gini (S)} - \sum_{i=1}^n \frac{|S_i|}{|S|} \text{Gini (S}_i) \quad (8)$$

Gain Informasi dengan Classification Error C Nilai Error diperoleh dari nilai atribut terkecil pada label kelas.

$$\text{C Error (S)} = 1 - \max_i [p(i|s)] \quad (9)$$

$$\text{Gain (S,A)} = \text{Error (S)} - \sum_{i=1}^n \frac{|S_i|}{|S|} \text{Error (S}_i) \quad (10)$$

### 2.2.3 Performance Evaluation

Model analysis model algoritma RF dan DT yang akan dilihat dari penelitian ini meliputi Evaluasi Model (EM), *Classification Accuracy* (CA), F1-Score, *Mattews Correlation Coefficient* (MCC), Precision (Prec), *Area Under the Curve*.

## 3 Hasil dan Pembahasan

### 3.1. Karakteristik Individu

Berdasarkan hasil analisis univariat yang ditunjukkan pada tabel 1 didapatkan ibu yang memiliki paritas > 4 sebanyak 102 orang (20,4%), jarak kehamilan ibu yang ≤ 2 tahun sebanyak 310 orang (62%), ibu pasca bersalin yang mengalami anemia sebanyak 124 orang (24,8%), ibu yang melahirkan bayi makrosomia sebanyak 60 orang (12 %), ibu yang mengalami komplikasi persalinan sebanyak 229 orang (45,8 %), ibu yang mengalami kehamilan ganda sebanyak 16 orang (3,2%), umur ibu yang berisiko sebanyak 132 orang (26,4%). Untuk lebih lengkapnya dapat dilihat pada resume tabel 1 berikut ini:

Tabel 1 Karakteristik Ibu Bersalin

No	Variabel & Kategori	Frekuensi	Persentase (%)
1	Paritas		
	> 4	102	20.4
	≤ 4	398	79.6
	Jumlah	500	100
2	Status Anemia		
	Ya	124	24.8
	Tidak	376	75.2
	Jumlah	500	100
3	Jarak Kehamilan		
	≤ 2	310	62
	> 2	190	38
	Jumlah	500	100
4	Bayi Makrosomia		
	> 4000 gram	60	12
	≤ 4000 gram	440	88
	Jumlah	500	100
5	Komplikasi Persalinan		
	Ya	229	45.8
	Tidak	271	54.2
	Jumlah	500	100
6	Umur		
	Berisiko	132	26.4
	Tidak Berisiko	368	73.6
	Jumlah	500	100
7	Kehamilan Ganda		
	>1	16	3.2
	1	484	96.8
	Jumlah	500	100

Pada tabel tersebut terlihat adanya variabel yang homogen yaitu variabel independen yang salah satu kategorinya mempunyai nilai < 15% yaitu variabel bayi makrosomia > 4000 gram sebanyak 60 orang (12%) dan variabel kehamilan ganda >1 sebanyak 16 orang (3,2 %).

### 3.2 Implementasi Decision Tree dan Random Forest

Peneliti membandingkan algoritma DT dan RF untuk mengukur evaluasi model, klasifikasi akurasi, F1-Score, Recall Model, MCC, Precision dan AUC. Hasil analisis didapatkan tabel 2 dibawah ini :

Tabel 2 Perbandingan Evaluasi Metrik antara Algoritma Decision Tree dan Random Forest

Algoritma	EM	CA	F1-Score	RM	MCC	Prec	AUC
Decision Tree	70 %	0.820	0.817	0.82	0.70	0.81	0.71
Random Forest	76 %	0.830	0.821	0.83	0.72	0.71	0.74

Pada tabel 2 evaluasi perbandingan model DT dan RF merupakan hasil dari cross-validation menggunakan model menggunakan 5-fold cross validation menunjukan algoritma RF menghasilkan evaluasi terbaik 76 %. Selanjutnya pada tingkat akurasi menunjukan Random RF memiliki akurasi yang terbaik sebesar 0.830 dibandingkan DT sebesar 0.817. Pengukuran model RF menunjukan terbaik dengan nilai 0.821 dibandingkan dengan algoritma DT sebesar 0.817, Hasil analisis menyatakan Random Forest menunjukan kemampuan lebih baik dengan nilai 0.830 dibandingkan algoritma Decision Tree. Evaluasi ini mengukur kualitas keseluruhan dari klasifikasi biner atau multiclass. Hasil ini menunjukkan RF memiliki nilai terbaik yaitu 0.724 dibandingkan dengan algoritma DT. Presisi model RF menunjukkan kemampuan yang lebih baik dengan nilai 0.816 dibandingkan dengan algoritma DT. Sedangkan pada AUC Evaluasi ini mengukur keseluruhan dari kurva ROC (Receiver Operating Characteristic) model yang menunjukan kualitas RF lebih baik sekitar 0.744.

Pada hasil penelitian dapat disimpulkan bahwa prediksi PPH dengan menggunakan algoritma RF memiliki tingkat akurasi lebih baik daripada algoritma DF. Menurut Deng et al (2019) [27] Random Forest merupakan pembelajaran ensemble algoritma dengan menggabungkan beberapa pengklasifikasi lemah, hasil akhirnya dipilih atau dirata-rata, sehingga hasilnya akan memiliki akurasi dan generalisasi yang lebih tinggi. Hal ini sesuai dengan penelitian dari Zhang et al, (2020) [28] bahwa Prediksi PPH menunjukan perbandingan hasil prediksi DIC dengan kumpulan data pengujian sebenarnya. Hal ini menunjukkan *Esembling Learning* dan Random Forest mencapai hasil prediksi yang lebih baik. Penelitian di Iran juga menggunakan pendekatan machine learning untuk memprediksi kejadian PPP pada 153 perempuan dengan diagnose PPP menyebutkan algoritma random forest dan decision tree termasuk dalam 5 algoritma yang memiliki performance terbaik dengan ROC\_AUC 0.86 % pada algoritma random forest dan 0.64 pada algoritma decision tree [29]. Pada penelitian lain menurut Iwendi, (2020) [30] menunjukan model yang mengimplementasikan algoritma Random Forest yang ditingkatkan oleh algoritma AdaBoost, dengan F1 Score sebesar 0,86 pada dataset pasien COVID-19. Studi telah menemukan bahwa algoritma Boosted Random Forest memberikan prediksi yang akurat bahkan pada dataset yang tidak seimbang.

#### 4 Kesimpulan

Perdarahan Postpartum (PPH) merupakan penyebab kematian tertinggi pada ibu hamil terutama di negara berkembang. Salah satu upaya yang dapat dilakukan untuk memberikan pengobatan yang tepat pada ibu adalah dengan melakukan skrining yang didukung dengan pemanfaatan teknologi yang saat ini berkembang pesat, dengan pendekatan Machine Learning (ML). Pada hasil penelitian dapat disimpulkan bahwa prediksi PPH dengan menggunakan algoritma RF memiliki tingkat akurasi lebih baik daripada algoritma DF Untuk meningkatkan efektivitas prediksi model ML dalam upaya meningkatkan pelayanan kesehatan untuk mengurangi prevalensi kejadian PPP perlu upaya personalisasi untuk setiap pasien hamil, serta untuk meningkatkan alokasi sumber daya dan mengurangi angka kematian serta kesakitan ibu, kita perlu melakukan studi masa depan dengan dataset yang lebih besar dan lebih banyak kasus PPH.

#### Ucapan Terimakasih

Terima kasih kepada Universitas Sari Mulia yang telah mendanai penelitian ini.

#### Daftar Rujukan

- [1] L. Simanjuntak, "Perdarahan Postpartum (Perdarahan Paskasalin)," *J. Visi Eksakta*, vol. 1, no. 1, pp. 1–10, 2020, doi: 10.51622/eksakta.v1i1.51.
- [2] A. Sunder *et al.*, "Maternal Obesity: An Obstetric Risk," *Cureus*, vol. 14, no. 9, pp. 1–9, 2022, doi: 10.7759/cureus.29345.
- [3] K. C. Wormer, R. T. Jamil, and S. B. Bryant, "Acute Postpartum Hemorrhage," *StatPearls*, May 2022.
- [4] WHO, *WHO recommendations: Uterotonics for the prevention of postpartum haemorrhage*. 2018.
- [5] Kemenkes, *profil kesehatan Indonesia 2019*. 2019.
- [6] Kemenkes, *Profil Kesehatan Indonesia Tahun 2020*. 2020.
- [7] Kemenkes RI., *Profil Kesehatan Indonesia*. 2021.
- [8] B. Boujarzadeh, A. Ranjbar, F. Banihashemi, V. Mehrnoush, F. Darsareh, and M. Saffari, "Machine learning approach to predict postpartum haemorrhage: A systematic review protocol," *BMJ Open*, vol. 13, no. 1, pp. 1–4, 2023, doi: 10.1136/bmjopen-2022-067661.
- [9] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert, and L. Cilar, "Interpretability of machine learning-based prediction models in healthcare," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 10, no. 5, p. e1379, Sep. 2020, doi: 10.1002/WIDM.1379.
- [10] W. Febriani, G. W. Nurcahyo, and S. Sumijan, "Diagnosa Penyakit Rubella Menggunakan Metode Fuzzy Tsukamoto," *J. Sistim Inf. dan Teknol.*, 2019, doi: 10.35134/jsisfotek.v1i3.4.
- [11] S. E. Dunkerton, Y. B. Jeve, N. Walkinshaw, E. Breslin, and T. Singhal, "Predicting Postpartum Hemorrhage (PPH) during Cesarean Delivery Using the Leicester PPH Predict Tool: A Retrospective Cohort Study," *Am. J. Perinatol.*, vol. 35, no. 2, pp. 163–169, 2018, doi: 10.1055/s-0037-1606332.
- [12] J. Liu *et al.*, "Machine learning-based prediction of postpartum hemorrhage after vaginal delivery: combining bleeding high risk factors and uterine contraction curve," *Arch. Gynecol. Obstet.*, vol. 306, no. 4, pp. 1015–1025, 2022, doi: 10.1007/s00404-021-06377-0.
- [13] J. Kang *et al.*, "Prediction model for massive transfusion in placenta previa during cesarean section," *Yonsei Med. J.*, vol. 61, no. 2, pp. 154–160, 2020, doi: 10.3349/ymj.2020.61.2.154.
- [14] nur andi fahira, abd rahman, and herman kurniawan, "12622-39746-1-Pb," *J. Kesehat. Tadulako*, vol. 5, no. 1, pp. 26–31, 2019.
- [15] A. Kristianingsih, H. Mukhlis, and E. Ermawati, "Faktor-faktor yang berhubungan dengan kejadian perdarahan postpartum di RSUD Pringsewu," *J. Wellnes*, vol. 2, no. February, pp. 309–313, 2020.
- [16] N. M. Abdulkareem and A. M. Abdulazeez, "Machine Learning Classification Based on Radom Forest Algorithm: A Review," *Int. J. Sci. Bus.*, vol. 5, no. 2, pp. 128–142, 2021, doi: 10.5281/zenodo.4471118.
- [17] U. Erdiansyah, A. Irmansyah Lubis, and K. Erwansyah, "Komparasi Metode K-Nearest Neighbor dan Random Forest Dalam Prediksi Akurasi Klasifikasi Pengobatan Penyakit Kulit," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 208, 2022, doi: 10.30865/mib.v6i1.3373.
- [18] S. Oktarian, S. Defit, and Sumijan, "Clustering Students' Interest Determination in School Selection Using the K-Means Clustering Algorithm Method," *J. Inf. dan Teknol.*, vol. 2, pp. 68–75, 2020, doi: 10.37034/jidt.v2i3.65.
- [19] A. Raza, H. U. R. Siddiqui, K. Munir, M. Almutairi, F. Rustam, and I. Ashraf, "Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction," *PLoS One*, vol. 17, no. 11 November, 2022, doi: 10.1371/journal.pone.0276525.
- [20] S. Syafwandi, D. Setyo Sembodo, A. Tua Munthe, and A. Sumarno, "Analysis of The Use of Sawdust Waste As Concrete Mixture Add Material Against Workability and Compressive Strength Concrete With Three Concrete Treatment Methods," *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 2, 2021, doi: 10.52088/ijesty.v1i2.109.

- [21] K. K. Venkatesh *et al.*, “Machine Learning and Statistical Models to Predict Postpartum Hemorrhage,” *Obstet. Gynecol.*, vol. 135, no. 4, pp. 935–944, 2020, doi: 10.1097/AOG.0000000000003759.
- [22] M. Schonlau and R. Y. Zou, “The random forest algorithm for statistical learning,” *Stata J.*, vol. 20, no. 1, pp. 3–29, 2020, doi: 10.1177/1536867X20909688.
- [23] P. Probst and A. L. Boulesteix, “To tune or not to tune the number of trees in random forest,” *J. Mach. Learn. Res.*, vol. 18, pp. 1–8, 2018.
- [24] D. Maisa Putra, O. Oktamianiza, M. Yuniar, and W. Fadhila, “Study Literature Review On Returning Medical Record Documents Using HOT-FIT Method,” *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 1, 2021, doi: 10.52088/ijesty.v1i1.102.
- [25] F. Wu, X. Liu, Y. Wang, X. Li, and M. Zhou, “Research on Evaluation Model of Hospital Informatization Level Based on Decision Tree Algorithm,” *Secur. Commun. Networks*, vol. 2022, 2022, doi: 10.1155/2022/3777474.
- [26] B. Charbuty and A. Abdulazeez, “Classification Based on Decision Tree Algorithm for Machine Learning,” *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, 2021, doi: 10.38094/jastt20165.
- [27] W. Deng, Y. Guo, J. Liu, Y. Li, D. Liu, and L. Zhu, “A missing power data filling method based on improved random forest algorithm,” *Chinese J. Electr. Eng.*, vol. 5, no. 4, pp. 33–39, 2019, doi: 10.23919/CJEE.2019.000025.
- [28] Y. Zhang, X. Wang, N. Han, and R. Zhao, “Ensemble Learning Based Postpartum Hemorrhage Diagnosis for 5G Remote Healthcare,” *IEEE Access*, vol. 9, pp. 18538–18548, 2021, doi: 10.1109/ACCESS.2021.3051215.
- [29] V. Mehrnoush, A. Ranjbar, M. V. Farashah, F. Darsareh, M. Shekari, and M. S. Jahromi, “Prediction of postpartum hemorrhage using traditional statistical analysis and a machine learning approach,” *AJOG Glob. Reports*, vol. 3, no. 2, p. 100185, 2023, doi: 10.1016/j.xagr.2023.100185.
- [30] C. Iwendi *et al.*, “COVID-19 patient health prediction using boosted random forest algorithm,” *Front. Public Heal.*, vol. 8, no. July, pp. 1–9, 2020, doi: 10.3389/fpubh.2020.00357.