



Predicting Outpatient Service Waiting Times with Random Forest Algorithm

Prediksi Waktu Tunggu Pelayanan Pasien Rawat Jalan dengan Algoritma Random Forest

**Rahayu Putri Munggaran¹, Mieke Nurmalasari^{2*},
Hosizah Hosizah³, Dewi Krismawati⁴**

^{1,2,3}Program Studi Manajemen Informasi Kesehatan, Fakultas Ilmu-Ilmu Kesehatan,
Universitas Esa Unggul, Indonesia

⁴Direktorat Analisis dan Pengembangan Statistik, Badan Pusat Statistik (BPS), Indonesia

E-Mail: ¹rahayuput1927@gmail.com, ²mieke@esaunggul.ac.id,
³hosizah@esaunggul.ac.id, ⁴dewikrisma@bps.go.id

Received Jul 11th 2024; Revised Aug 25th 2024; Accepted Sep 29th 2024; Available Online Nov 24th 2024

Corresponding Author: Mieke Nurmalasari

Copyright © 2025 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

Waiting time for services is a crucial step patients go through to receive healthcare, starting from registration to the doctor's examination. This study aims to predict the waiting time for outpatient services using the Random Forest algorithm at Dr. Soeharto Heerdjan Mental Hospital. The prediction is expected to facilitate staff work and can be integrated into an online application to reduce patient congestion. Data mining methods were applied using Orange Data Mining software with the Random Forest algorithm. The study was conducted using 2,109 data points collected over three months in 2023, which after preprocessing resulted in 1,508 data points with 8 attributes: age, gender, clinic, selected service, arrival time, time before meeting the doctor, waiting duration, health insurance, and patient category. The data were divided into training data (1,055) and testing data (452). The prediction results showed high accuracy with AUC 98.2%, CA 97.6%, F1 97.6%, precision 97.6%, and recall 97.3%. The ROC-curve model could distinguish three categories of waiting time: fast, slow, and normal, with accuracy values close to 1. Visualization using the Pythagorean Forest helped identify the category or pattern of waiting time with high accuracy.

Keyword: Classification, Data Mining, Outpatient Services, Prediction, Random Forest, Waiting Time Prediction

Abstrak

Waktu tunggu pelayanan merupakan salah satu langkah yang harus dilalui pasien untuk mendapatkan pelayanan kesehatan, dimulai dari pendaftaran hingga pemeriksaan oleh dokter. Penelitian ini bertujuan memprediksi waktu tunggu pelayanan pasien rawat jalan menggunakan algoritma Random Forest di Rumah Sakit Jiwa Dr. Soeharto Heerdjan. Prediksi ini diharapkan mempermudah pekerjaan petugas dan dapat diintegrasikan ke dalam aplikasi online untuk mengurangi penumpukan pasien. Metode data mining diterapkan menggunakan aplikasi Orange Data Mining dengan algoritma Random Forest. Penelitian dilakukan menggunakan 2.109 data dari tiga bulan di tahun 2023, yang setelah preprosesing menghasilkan 1.508 data dengan 8 atribut: usia, jenis kelamin, poliklinik, layanan yang dipilih, waktu datang, waktu sebelum bertemu dokter, durasi waktu tunggu, jaminan kesehatan, dan kategori pasien. Data dibagi menjadi dua bagian, yaitu data training sebanyak 1.055 dan data testing sebanyak 452. Hasil prediksi menunjukkan akurasi tinggi dengan nilai AUC 98,2%, CA 97,6%, F1 97,6%, precision 97,6%, dan recall 97,3%. Model ROC-curve dapat memisahkan tiga kategori waktu tunggu yaitu cepat, lambat, dan normal, dengan nilai akurasi mendekati 1. Visualisasi menggunakan Pythagorean Forest membantu mengidentifikasi kategori atau pola waktu tunggu pasien dengan akurasi yang tinggi.

Kata Kunci: Data Mining, Klasifikasi, Prediksi Waktu Tunggu, Random Forest, Rawat Jalan

1. PENDAHULUAN

Waktu tunggu pasien adalah salah satu komponen yang dapat menyebabkan ketidakpuasan pasien. Lamanya waktu tunggu dapat mencerminkan kemampuan rumah sakit dalam mengelola pelayanan sesuai

dengan harapan pasien. Indikator waktu tunggu pelayanan pasien rawat jalan di rumah sakit adalah 60 menit, sesuai dengan Kemenkes Nomor 129/Menkes/SK/II/2008. Jika waktu tunggu melebihi 60 menit, pelayanan tersebut dianggap tidak memenuhi standar yang berlaku.

Penelitian ini menemukan bahwa waktu tunggu pasien rawat jalan dari 10 pasien menunjukkan 70% pasien memiliki waktu tunggu lebih lama dari standar, yaitu sekitar 56-75 menit [1]. Waktu tunggu yang lama ini juga ditemukan dalam pelayanan di Rumah Sakit Umum Daerah (RSUD) Kabupaten Karanganyar. Terdapat beberapa keluhan dari kotak saran mengenai antrian pasien yang sangat lama di bagian pelayanan rawat jalan. Capaian indikator waktu tunggu di rawat jalan sebesar 55% dari standar yaitu 60%. Hal ini menunjukkan rumah sakit belum memenuhi indikator waktu tunggu di rawat jalan [2].

Penelitian lain mengenai kualitas pelayanan program posyandu jiwa menunjukkan bahwa pasien dengan gangguan jiwa yang harus menunggu lama bisa menjadi gelisah dan bahkan mengamuk [3]. Pengaruh Waktu tunggu juga berdampak kurangnya terkarakterisasi dengan baik untuk perawatan primer, perawatan spesialis, janji temu dengan dokter spesialis, dan waktu yang dihabiskan di ruang tunggu [4]. Permasalahan dapat difinimalisir menggunakan metode data mining seperti Classification and Regression Tree (CART) dan Random Forest mengambil sampel 10 pasien pertama dan menemukan rata-rata waktu tunggu adalah 26 menit berdasarkan umur, jenis kelamin, dan poliklinik yang dipilih pasien [5].

Tujuan penelitian ini adalah melakukan seleksi data, prediksi, dan evaluasi keakuratan kinerja model *machine learning*. Penelitian ini memprediksi waktu tunggu menggunakan algoritma pembelajaran mesin atau data mining. Prediksi memperlihatkan kejadian yang terjadi pada keadaan tertentu. Waktu tunggu pasien merupakan komponen yang penting [6]. Prediksi menggunakan data mining metode tepat untuk memprediksi dengan data yang besar dan kompleks. Algoritma *Random Forest* digunakan karena mampu menghasilkan prediksi dengan akurasi tinggi, eror yang relatif kecil, serta mudah dipahami [7]. Penerapan data mining ini diharapkan dapat mempermudah rumah sakit dalam menentukan waktu tunggu rawat jalan. Implementasi data mining pada data waktu tunggu pasien rawat jalan diharapkan mampu memprediksi durasi waktu tunggu, yang kemudian dapat diterapkan dalam aplikasi pendaftaran online. Dengan demikian, pasien dapat melihat perkiraan waktu tunggu saat memilih poliklinik dan layanan yang diinginkan dalam aplikasi tersebut.

2. BAHAN DAN METODE

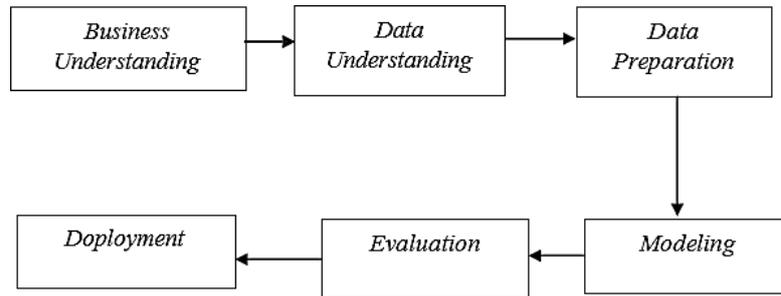
Penelitian ini menggunakan metode data mining dengan algoritma *Random Forest* dengan menggunakan *Orange Data Mining*. *Random Forest* bekerja dengan mengambil sampel acak dari dataset, membentuk *decision tree* pada setiap sampel, mendapatkan hasil prediksi dari setiap tree, dan melakukan voting pada hasil prediksi untuk menghasilkan prediksi akhir [8].

Kategori waktu tunggu pasien pada penelitian ini dibagi menjadi tiga yaitu lambat (<60 menit), normal (60 menit), dan cepat (>60 menit) sesuai dengan standar Kemenkes Nomor 129/Menkes/SK/II/2008 [9]. Penelitian ini juga mengacu pada penelitian Akbar [10], yang menggunakan teori Fetter dan Thompson [11], di mana waktu tunggu dibagi menjadi *First Waiting Time* (waktu dari pasien datang hingga mendaftar), *True Waiting Time* (waktu dari pendaftaran hingga dipanggil dokter), dan *Total Primary Waiting Time* (gabungan dari seluruh waktu tunggu).

Penelitian mengenai waktu tunggu pelayanan rawat jalan telah banyak dikaji dalam berbagai studi sebelumnya. Penelitian Fajrin, Haeruddin, dan Ahri menyelidiki waktu tunggu pelayanan dengan memperhitungkan faktor-faktor seperti umur, jenis kelamin, jenis pasien, poliklinik tujuan, dan jaminan kesehatan yang dipilih pasien [1]. Sementara itu penelitian yang dilakukan oleh Handayani, Mustafid, dan Surarso menggunakan atribut seperti umur, jenis kelamin, nama pelayanan, serta waktu mulai dan selesai pelayanan. Penelitian ini memfokuskan pada faktor-faktor usia, jenis kelamin, jaminan kesehatan, kategori pasien, poliklinik, dan jenis pelayanan yang dipilih oleh pasien dalam menganalisis waktu tunggu pelayanan rawat jalan [12]. menggunakan atribut umur, jenis kelamin, nama pelayanan, waktu pasien mulai, waktu pasien selesai. Penelitian ini menggunakan usia, jenis kelamin, jaminan kesehatan, kategori pasien, poliklinik dan pelayanan yang dipilih oleh pasien.

Penelitian ini adalah analisis prediktif yang menggunakan data sekunder rekam medis pasien rawat jalan sebanyak 2.109 data dari poliklinik psikiatri dewasa, anak dan remaja, serta poliklinik geriatri. Variabel yang diamati meliputi usia, jenis kelamin, jaminan kesehatan, kategori kedatangan pasien, nama poliklinik, nama layanan, waktu datang, waktu sebelum masuk ke dokter, dan durasi waktu tunggu. *Random Forest* adalah pengembangan dari metode *Decision Tree* yang menggunakan beberapa *Decision Tree*. Setiap *Decision Tree* dilatih menggunakan sampel individu, dan atribut yang digunakan untuk pemisahan pada pohon dipilih secara acak dari subset atribut [13]. *Random Forest* memiliki beberapa kelebihan, yaitu dapat meningkatkan hasil akurasi jika terdapat data yang hilang, dan untuk *resisting outliers*, serta efisien untuk penyimpanan sebuah data [14]. *Random Forest* memiliki performa prediksi yang baik dalam regresi serta klasifikasi [15]. *Random forest* juga menunjukkan performa yang lebih baik saat disandingkan dengan metode lain 9 seperti, *partial least squares regression*, *support vector machine* dan *neural network*. Seleksi fitur tentu *Random Forest* dapat bekerja pada big data dengan parameter yang kompleks secara efektif [16]. Teknik sampling yang digunakan adalah *purposive sampling* dan analisis data dilakukan dengan metode CRISP-DM. CRISP DM sebagai salah

satu standard dalam data mining dan sebagai salah satu bentuk model penambangan data yang paling banyak digunakan karena kelebihanannya yang memecahkan masalah yang ada di industri penambangan data [17]. CRISP DM terdiri dari enam tahap seperti Gambar 1.



Gambar 1. Tahapan CRISP-DM

Metode ini menginterpretasikan berbagai permasalahan ke dalam teknik data mining sebagai berikut:

1. **Pemahaman Bisnis (*Business Understanding*)**
Memahami masalah penelitian, yaitu prediksi waktu tunggu pasien untuk meningkatkan efisiensi dan kepuasan.
2. **Pemahaman Data (*Data Understanding*)**
Mengumpulkan data rekam medis elektronik yang terdiri dari 9 atribut: usia, jenis kelamin, poliklinik, layanan yang dipilih, waktu datang, waktu sebelum bertemu dokter, durasi waktu tunggu, jaminan kesehatan, dan kategori pasien.
3. **Persiapan Data (*Data Preparation*)**
Melakukan pre-processing untuk membersihkan dan mempersiapkan data mentah.
4. **Pemodelan (*Modelling*)**
Membagi dataset menjadi data training dan testing, kemudian menggunakan algoritma Random Forest untuk pemodelan.
5. **Evaluasi (*Evaluation*)**
Menguji kinerja model dengan confusion matrix.
6. **Perencanaan (*Deployment*)**
Menerapkan model yang telah dievaluasi, dengan menghasilkan nilai model, skor keputusan, dan interpretasi hasil untuk implementasi praktis, seperti aplikasi pendaftaran online [18].

3. HASIL DAN PEMBAHASAN

3.1. Seleksi Data Terhadap Variabel Waktu Tunggu Pelayanan Pasien Rawat Jalan Menggunakan Algoritma Random Forest

Seleksi data dilakukan untuk menentukan variabel yang akan digunakan. Penelitian ini menggunakan data sebanyak 2.109 data pasien rawat jalan. Tahap preprocessing dilakukan untuk menghilangkan beberapa permasalahan setelah dilakukan tahap preprocessing data menjadi 1.507 data dengan 8 atribut seperti Tabel 1.

Tabel 1. Data Waktu Tunggu Pasien Rawat Jalan

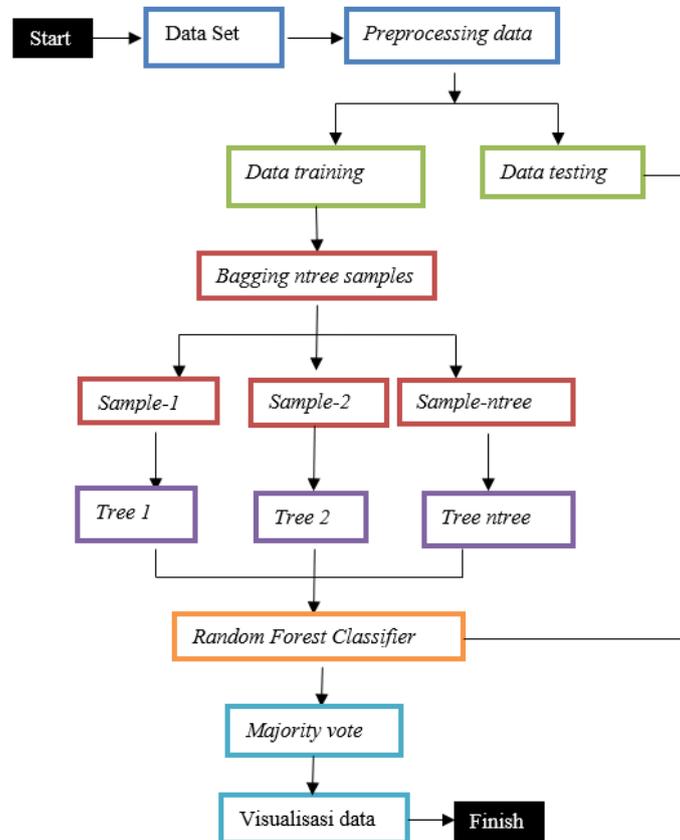
Umur	Jenis kelamin	Status kepesertaan	Kategori pasien	Poliklinik	Pelayanan	Waktu Kedatangan	Durasi Waktu Tunggu
7	P	BPJS	Pasien lama	Anak	Neuroleptic Therapy	7	30 menit
25	L	BPJS	Pasien lama	Geriatric	Counselling	8	60 menit
50	L	Umum	Pasien baru	Dewasa	Educational Therapy	10	40 menit

Tabel 1 menunjukkan data waktu tunggu pasien rawat jalan yang mencakup beberapa atribut kunci. Mayoritas pasien dalam data ini adalah usia 20-59 tahun (80,8%) dan jenis kelamin laki-laki (62,7%), dengan sebagian besar menggunakan BPJS sebagai jaminan kesehatan (94,1%). Pasien lama mendominasi (96,5%), sering kali mengunjungi poliklinik psikiatri dewasa (81,8%) untuk pelayanan seperti neuroleptic therapy

(81,4%). Waktu kedatangan utama adalah pada pukul 8 pagi (23,0%), sementara durasi waktu tunggu yang paling umum adalah antara 61 hingga 100 menit (51,9%).

3.2. Prediksi Waktu Tunggu Pelayanan Pasien Rawat Jalan Menggunakan Metode Random Forest

Pada aplikasi *orange data mining*, dataset dibagi menjadi dua yaitu 70% *data training* dan *data testing* 30%. [19]. Pemodelan dilakukan untuk merancang data mining dalam memprediksi waktu tunggu pelayanan menggunakan alur seperti Gambar 2.



Gambar 2. Flowchat Proses Random Forest

3.3. Evaluasi Keakuratan Metode Random Forest dalam Memprediksi Waktu Tunggu Pelayanan Pasien Rawat Jalan

Evaluasi hasil pengujian data waktu tunggu pelayanan pasien rawat jalan dengan menggunakan prediksi random forest menghasilkan test and score seperti pada Tabel 2.

Tabel 2. Hasil Prediksi *Test and Score*

Model	AUC	CA	F1	Precision	Recall
Random Forest	0.982	0.976	0.976	0.976	0.976

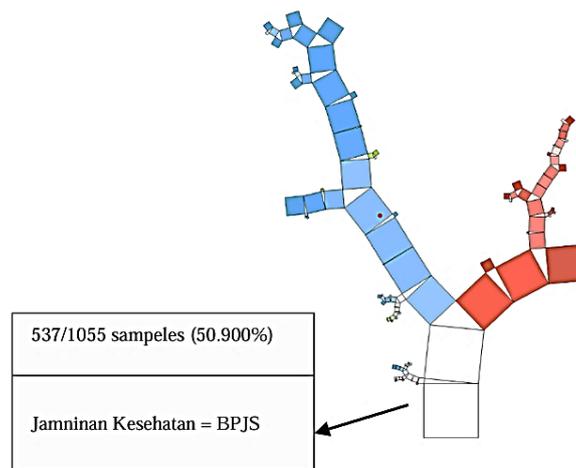
Hasil evaluasi menunjukkan bahwa model *Random Forest* memiliki nilai *Area Under Curve* (AUC) sebesar 0,982 (98,2%). AUC menggambarkan probabilitas bahwa model dapat membedakan antara kelas positif dan negatif dengan baik. Semakin besar nilai AUC, semakin baik performa model dalam melakukan klasifikasi. Nilai *Classification Accuracy* (CA) sebesar 0,976 (97,6%) menunjukkan persentase prediksi yang benar dari keseluruhan prediksi. Nilai F1 sebesar 0,976 (97,6%) merupakan rata-rata harmonis antara *Precision* dan *Recall*, yang memberikan keseimbangan antara keduanya. *Precision* sebesar 0,976 (97,6%) menunjukkan proporsi prediksi positif yang benar dari keseluruhan hasil prediksi positif. *Recall* sebesar 0,976 (97,6%) menunjukkan kemampuan model dalam mendeteksi semua data positif dari keseluruhan data positif.

Confusion matrix pada Tabel 3 menunjukkan jumlah *instance* antara kelas prediksi dengan kelas aktual. Hasilnya, prediksi yang benar atau sesuai adalah 442 (97,6%), sedangkan yang tidak sesuai adalah 11 (2,4%). *Confusion matrix* ini juga menunjukkan pembagian *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Akurasi model juga dapat dievaluasi melalui grafik *Receiver Operating Characteristics* (ROC).

Tabel 3. Confusion Matrix Random Forest

	Cepat	Lambat	Normal	
Cepat	182	6	0	188
Lambat	1	214	4	219
Normal	0	0	45	45
	183	220	49	452

ROC adalah grafik yang menunjukkan hubungan antara *True Positive Rate* (sumbu Y) dan *False Positive Rate* (sumbu X) untuk menilai seberapa baik model. Dalam prediksi ini, waktu tunggu pelayanan dibagi menjadi tiga kategori: cepat, lambat, dan normal. Kategori cepat memiliki nilai ROC sebesar 0,490, diikuti dengan kategori lambat memiliki nilai 0,510, dan kategori normal memiliki nilai 0,557. Evaluasi ini menunjukkan bahwa model *Random Forest* mampu memberikan prediksi yang akurat dan andal dalam memprediksi waktu tunggu pasien rawat jalan, yang dapat digunakan untuk meningkatkan efisiensi pelayanan di rumah sakit. Visualisasi *Pythagorean forest* dan *Pythagorean tree* digunakan agar dapat memberikan informasi. Penelitian ini menunjukkan kotak paling besar menampilkan akar [20], dimana ini menjadi faktor paling berpengaruh pada waktu tunggu pelayanan rawat jalan seperti pada Gambar 3.



Gambar 3. Pythagoras Tree Waktu Tunggu Pelayanan

Gambar 3 menunjukkan Visualisasi *Pythagorean forest* dan *Pythagorean tree* digunakan agar dapat memberikan informasi. Penelitian ini menunjukkan kotak paling besar menampilkan akar. Atribut jaminan Kesehatan menjadi akar awal di pohon *Pythagoras* yang artinya Jaminan Kesehatan menjadi faktor paling berpengaruh pada waktu tunggu pelayanan rawat jalan. Hasil seleksi data melalui *preprocessing* menunjukkan bahwa dari 2.109 data pasien rawat jalan yang digunakan, sebanyak 602 data dengan *missing values*, sehingga data yang digunakan untuk memprediksi dengan *data mining* berjumlah 1.507.

Hasil data yang sudah di preprosesing sebanyak 1.057 dibagi menjadi dua yaitu *data training* 1.055 dan *data testing* sebanyak 452. Hasil akurasi prediksi 452 data menggunakan nilai *test and score* menunjukkan nilai AUC, CA, F1, Precision dan recall semua nilai mendekati angka 1 yang artinya hasil positif dan benar. Confusion matrix 3 class digunakan untuk melihat hasil evaluasi prediksi *true positif*, *true negative*, *false positif* dan *false negative* kategori cepat diprediksi cepat sebanyak 182. Normal diprediksi normal sebanyak 45 dan lambat diprediksi akan lambat sebanyak 214 data. Data yang tidak sesuai didapatkan sebanyak 11 data.

4. KESIMPULAN

Penelitian ini berhasil melakukan selection dan preprocessing data, mengidentifikasi dan menangani data yang hilang, sehingga 1.507 data dapat diolah. Implementasi algoritma Random Forest menunjukkan hasil akurasi yang sangat tinggi, dengan nilai AUC, CA, F1, Precision, dan Recall mendekati 1, yang berarti hasil prediksi sangat akurat. Confusion matrix tiga kelas digunakan untuk mengevaluasi prediksi, menunjukkan bahwa kategori cepat diprediksi dengan benar sebanyak 182 kali, kategori normal sebanyak 45 kali, dan kategori lambat sebanyak 214 kali, dengan hanya 11 prediksi yang tidak sesuai. Evaluasi waktu menggunakan visualisasi Pythagorean Tree menunjukkan bahwa atribut jaminan kesehatan (BPJS) merupakan faktor yang signifikan dalam model prediksi. Hasil ini mengindikasikan bahwa Random Forest adalah metode yang efektif untuk memprediksi waktu tunggu pasien rawat jalan, dan dapat digunakan untuk meningkatkan efisiensi pelayanan di rumah sakit.

REFERENSI

- [1] K. Fajrin, Haeruddin, and R. A. Ahri, "Faktor Yang Mempengaruhi Waktu Tunggu Pelayanan Rawat Jalan Pasien di RSUD Kota Makasar," *Wind. Public Heal. J.*, vol. 2, no. 5, pp. 827–835, 2021.
- [2] Oktavy Budi Kusumawardhani, Rada Febria Kurniawati, "Faktor-Faktor Yang Mempengaruhi Petugas dalam Pelayanan Waktu Tunggu Pasien di Rumah Sakit," no. 44, pp. 125–136, 2023.
- [3] S. Dwieka Septiani AdiwinoIlmu, A. Negara, F. Ilmu, and U. N. Surabaya, "Kualitas Pelayanan Program Posyandu Kualitas Pelayanan Program Posyandu Jiwa terhadap Pasien ODGJ(Orang Dengan Gangguan Jiwa) di Desa Mojotamping," 2021
- [4] D. McIntyre and C. K. Chow, "Waiting Time as an Indicator for Health Services Under Strain: A Narrative Review," *Inq. (United States)*, vol. 57, 2020, doi: 10.1177/0046958020910305.
- [5] J. Chen, K. Li, Z. Tang, K. Bilal, and K. Li, "A parallel patient treatment time prediction algorithm and its applications in hospital queuing-recommendation in a big data environment," *IEEE Access*, vol. 4, pp. 1767–1783, 2016, doi: 10.1109/ACCESS.2016.2558199.
- [6] M. D. Wahyudi, "Penerapan Data Mining Dengan Algoritma C4. 5 Dalam Prediksi Penjualan Buku," *J. Teknorama (Informatika dan ...)*, vol. 1, no. 1, pp. 1–6, 2023,
- [7] A. Atalan, "Neural Network and Random Forest Algorithms for Estimation of the Waiting Times Based on the DES in ED," *Int. Conf. Contemp. Acad. Res.*, vol. 1, no. 1, pp. 14–20, 2023.
- [8] A. Cutler and D. R. Cutler, "Ensemble Machine Learning Random Forest," *Ensemble Mach. Learn.*, no. February 2014, 2012, doi: 10.1007/978-1-4419-9326-7.
- [9] "Keputusan Menteri Kesehatan Republik Indonesia No 377/Menkes/SK/III/2007." 2007.
- [10] M. R. Akbar, *Kepuasan Pasien Jaminan Kesehatan Nasional Terhadap Waktu Tunggu di Puskesmas dan Klinik Mitra BPJS*.
- [11] R. B. Fetter and J. D. Thompson, "Patients ' Waiting Time and Doctors ' Idle Time in the Outpatient _ Setting," 1996.
- [12] D. P. Handayani, M. Mustafid, and B. Surarso, "Patient Queue Systems in Hospital Using Patient Treatment Time Prediction Algorithm," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 4, pp. 45–54, 2020, doi: 10.22219/kinetik.v5i1.1001.
- [13] S. Trivedi and N. Patel, "The Impact of Artificial Intelligence Integration on Minimizing Patient Wait Time in Hospitals," *Res. Rev. Sci. Technol.*, vol. 3, no. 1, pp. 21–35, 2020,
- [14] M. Hanko *et al.*, "Random Forest–Based Prediction of Outcome and Mortality in Patients with Traumatic Brain Injury Undergoing Primary Decompressive Craniectomy," *World Neurosurg.*, vol. 148, pp. e450–e458, 2021, doi: 10.1016/j.wneu.2021.01.002.
- [15] P. Patil and S. Thakur, "Patient waiting time prediction in hospital queuing system using improved random forest in big data," *IEEE Int. Conf. Issues Challenges Intell. Comput. Tech. ICICT 2019*, 2019, doi: 10.1109/ICICT46931.2019.8977717.
- [16] J. Joseph, S. Senith, A. A. Kirubaraj, and J. S. R. Ramson, "Machine Learning for Prediction of Wait Times in Outpatient Clinic," *Procedia Comput. Sci.*, vol. 215, no. 2022, pp. 230–239, 2022, doi: 10.1016/j.procs.2022.12.026.
- [17] A. Azevedo and M. F. Santos, "KDD , SEMMA AND CRISP-DM : A PARALLEL OVERVIEW Ana Azevedo and M . F . Santos," *IADIS Eur. Conf. Data Min.*, pp. 182–185,
- [18] P. C. Ncr *et al.*, "Crisp-Dm," *SPSS inc*, vol. 78, pp. 1–78,
- [19] H. Hidayat, A. Sunyoto, and H. Al Fatta, "Klasifikasi Penyakit Jantung Menggunakan Random Forest Clasifier," *J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan)*, vol. 7, no. 1, pp. 31–40, 2023
- [20] E. W. Ambarsari and H. Herlinda, "Membangun Pythagoras Sebagai Visualisasi Random Forest Untuk Pemodelan Pohon Keputusan," *Fakt. Exacta*, vol. 13, no. 3, pp. 132–139, 2020.