

INTEGRASI MACHINE LEARNING DAN ANALISIS SPASIAL UNTUK PREDIKSI WILAYAH RAWAN TUBERKULOSIS DI PROVINSI SUMATERA UTARA

Fanny Ramadhani¹, Said Iskandar Al-Idrus², Dian Septiana³, Arnita⁴, Diah Retno Wahyuningrum⁵

1,2) Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan Indonesia

3,4) Statistik, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan Indonesia

5) Gizi, Fakultas Teknik, Universitas Negeri Medan Indonesia

Article Info	ABSTRACT
Article history: Received: 15 Juni 2025 Revised: 03 Juli 2025 Accepted: 07 Juli 2025	Abstrak Tuberkulosis (TBC) masih menjadi masalah kesehatan masyarakat yang signifikan di Provinsi Sumatera Utara. Tingginya prevalensi TBC, terutama di daerah padat penduduk dan terbatasnya akses terhadap layanan kesehatan, menjadi tantangan utama dalam upaya pengendalian penyakit ini. Penelitian ini bertujuan memprediksi wilayah rawan TBC dengan mengintegrasikan algoritma <i>machine learning</i> dan analisis spasial. Data sekunder dikumpulkan dari Sistem Informasi Tuberkulosis Nasional (SITB), Badan Pusat Statistik (BPS), dan <i>shapefile</i> administratif kabupaten/kota di Sumatera Utara. Variabel prediktor meliputi kepadatan penduduk, status gizi, jumlah fasilitas kesehatan, tingkat kemiskinan, kualitas hunian, dan cakupan imunisasi dasar lengkap. Model prediksi dikembangkan menggunakan algoritma Random Forest, sedangkan analisis spasial dilakukan dengan QGIS untuk menghasilkan peta risiko TBC. Model Random Forest menunjukkan performa yang sangat baik dengan akurasi sebesar 90,2%, precision 86,7%, recall 84,5%, dan F1-score 84,1%. Variabel yang paling berpengaruh terhadap kerawanan TBC adalah kepadatan penduduk, kualitas hunian, dan jumlah fasilitas kesehatan. Peta risiko mengidentifikasi Kota Medan, Kabupaten Deli Serdang, dan Kabupaten Labuhanbatu sebagai zona merah. Analisis spasial lebih lanjut menggunakan Moran's I (0,47; $p < 0,01$) dan Getis-Ord Gi* menunjukkan adanya pola kluster signifikan dan hotspot TBC di wilayah tersebut. Integrasi algoritma machine learning dan analisis spasial efektif dalam memprediksi dan memetakan wilayah rawan TBC di Sumatera Utara. Pendekatan ini dapat menjadi dasar perencanaan intervensi kesehatan berbasis bukti dan lokasi, serta direplikasi untuk penyakit menular lainnya yang memiliki karakteristik spasial serupa. Kata kunci: Tuberkulosis, Machine Learning, Random Forest, Spasial, Sumatera Utara Abstract <i>Tuberculosis (TB) remains a significant public health issue in North Sumatra Province. The high prevalence of TB, particularly in densely populated areas and regions with limited access to healthcare services, presents major challenges for disease control efforts. This study aims to predict TB-prone areas by integrating machine learning algorithms and spatial analysis. Secondary data were collected from the National Tuberculosis Information System (SITB), the Central Bureau of Statistics (BPS), and administrative shapefiles of districts/cities in North Sumatra. Predictor variables included population density, nutritional status, number of healthcare facilities, poverty rate, housing quality, and basic immunization coverage. The predictive model was developed using the Random Forest algorithm, while spatial analysis was conducted with QGIS to generate TB risk maps. The Random Forest model demonstrated strong performance with an accuracy of 90.2%, precision of 86.7%, recall of 84.5%, and an F1-score of 84.1%. The most influential variables affecting TB risk were population density, housing quality, and availability of healthcare facilities. The risk map identified Medan City, Deli Serdang District, and Labuhanbatu District as high-risk (red zone) areas. Further spatial analysis using Moran's I (0.47; $p < 0.01$) and Getis-Ord Gi* revealed significant clustering patterns and TB hotspots in these regions. The integration of machine learning algorithms and spatial analysis proved effective in predicting and mapping TB-prone areas in North Sumatra. This approach can serve</i>

as a foundation for evidence-based, location-specific health intervention planning and may be replicated for other infectious diseases with similar spatial characteristics.

Keywords: *Tuberculosis, Machine Learning, Random Forest, Spatial Analysis, North Sumatra.*

Djtechno: Jurnal Teknologi Informasi oleh Universitas Dharmawangsa Artikel ini bersifat open access yang didistribusikan di bawah syarat dan ketentuan dengan Lisensi Internasional Creative Commons Attribution NonCommercial ShareAlike 4.0 ([CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/)).



Corresponding Author:

E-mail : fannyr@unimed.ac.id

1. PENDAHULUAN

Tuberkulosis (TBC) adalah penyakit infeksi menular yang masih menjadi salah satu masalah kesehatan masyarakat utama di Indonesia, termasuk di Provinsi Sumatera Utara [1]. Meskipun telah ada berbagai upaya pengendalian, angka kejadian TBC di beberapa wilayah tetap tinggi dan cenderung tidak merata. Tuberkulosis (TBC) merupakan penyakit menular yang disebabkan oleh bakteri *Mycobacterium tuberculosis*, dan masih menjadi masalah kesehatan masyarakat yang besar di Indonesia [2]. Menurut WHO Global TB Report 2023 [3], Indonesia merupakan negara dengan beban TBC tertinggi kedua di dunia. Di Sumatera Utara, angka kasus TBC terus meningkat setiap tahun, terutama di wilayah urban dan semi-urban dengan kepadatan tinggi serta akses layanan kesehatan yang tidak merata [4]. Faktor-faktor seperti kepadatan penduduk, kondisi sosial ekonomi, kualitas lingkungan, dan akses terhadap layanan kesehatan menjadi determinan penting yang memengaruhi penyebaran penyakit ini [5]. Oleh karena itu, identifikasi wilayah rawan TBC secara tepat sangat diperlukan untuk mengarahkan intervensi yang efektif dan efisien.

Dalam beberapa tahun terakhir, perkembangan teknologi informasi dan kecerdasan buatan, khususnya machine learning, memberikan peluang baru dalam pengolahan data kesehatan yang kompleks [6][7]. Berbagai studi telah menunjukkan efektivitas penggunaan machine learning dalam prediksi penyakit menular, termasuk TBC [8]. Algoritma Random Forest dikenal memiliki performa yang baik dalam klasifikasi dan regresi, serta mampu menangani data dengan banyak variable [9]. Integrasi pendekatan spasial dan machine learning dalam penelitian kesehatan dapat membantu mengidentifikasi kluster penyakit dan memperkuat kebijakan berbasis data [10]. Namun, di Sumatera Utara, aplikasi pendekatan ini masih terbatas. Metode ini mampu

menganalisis berbagai variabel secara simultan untuk memprediksi risiko kejadian penyakit dengan tingkat akurasi yang lebih tinggi dibandingkan pendekatan tradisional [11]. Selain itu, analisis spasial menggunakan Geographic Information System (GIS) menjadi alat yang efektif untuk memvisualisasikan distribusi penyakit dan mengidentifikasi pola kluster spasial secara geografis [12].

Penelitian ini bertujuan mengintegrasikan metode machine learning dan analisis spasial guna memprediksi dan memetakan wilayah rawan TBC di Provinsi Sumatera Utara. Dengan pendekatan ini, diharapkan dapat memberikan informasi yang lebih akurat dan komprehensif bagi pembuat kebijakan dan tenaga kesehatan dalam merencanakan program pengendalian TBC yang lebih tepat sasaran.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif prediktif dengan mengintegrasikan algoritma machine learning dan analisis spasial untuk memetakan wilayah rawan Tuberkulosis (TBC) di Provinsi Sumatera Utara. Proses penelitian dilakukan secara sistematis melalui beberapa tahapan yang mencakup pengumpulan data, pengolahan data, pemodelan prediktif, dan visualisasi spasial [13].

1. Desain dan Lokasi Penelitian

Penelitian ini bersifat non-eksperimental dan berbasis data sekunder, dengan cakupan wilayah seluruh kabupaten/kota di Provinsi Sumatera Utara. Analisis dilakukan pada tingkat agregasi wilayah administratif.

2. Sumber dan Jenis Data

Data yang digunakan adalah data sekunder dari berbagai sumber resmi, yaitu:

- a. Data Kasus TBC: Diambil dari Sistem Informasi Tuberkulosis Nasional (SITB) Kementerian Kesehatan RI.
- b. Data Sosial-Demografi dan Lingkungan: Diambil dari Badan Pusat Statistik (BPS), meliputi data jumlah penduduk, fasilitas kesehatan, status gizi, kemiskinan, dan cakupan imunisasi.
- c. Data Spasial: Shapefile administrasi kabupaten/kota diperoleh dari Badan Informasi Geospasial (BIG).

3. Variabel Penelitian

- a. Variabel Target (Label): Klasifikasi wilayah rawan TBC berdasarkan kuartil tertinggi jumlah kasus TBC.
- b. Variabel Prediktor:
 - Kepadatan penduduk (jiwa/km²),
 - Status gizi buruk (persentase),
 - Jumlah fasilitas kesehatan (per wilayah),
 - Tingkat kemiskinan (persentase),
 - Kualitas hunian (persentase rumah tidak layak huni),
 - Cakupan imunisasi dasar lengkap (persentase).

4. Tahapan Penelitian

A. Pra-Pemrosesan Data [14] [15],

- a. Konsolidasi Dataset adalah menyatukan data numerik dari berbagai sumber berdasarkan ID wilayah.
- b. Pembersihan Data dengan cara menangani missing values dan inkonsistensi data.
- c. Normalisasi/Standardisasi: Dilakukan agar skala antar variabel menjadi sebanding.

B. Pemodelan Machine Learning

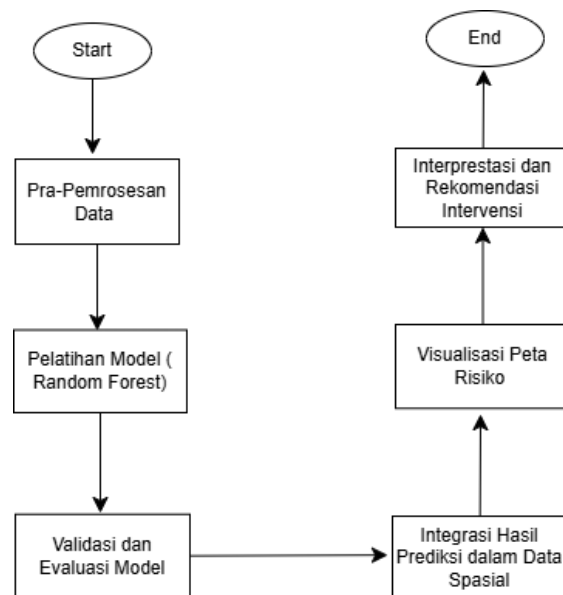
- a. Pemilihan Algoritma. Pada penelitian ini, Algoritma Random Forest dipilih karena kemampuannya menangani data dengan banyak variabel dan interpretabilitas yang tinggi.
- b. Pelatihan Model. Pada proses pelatihan model, Dataset dibagi menjadi data latih dan data uji (80:20). Kemudian Model dievaluasi menggunakan 5-fold cross-validation[8]
- c. Evaluasi Model. Pada penelitian ini, evaluasi model menggunakan Metrik evaluasi. Metrik evaluasi yang digunakan adalah akurasi, precision, recall, dan F1-score. Feature importance diekstraksi untuk mengetahui variabel paling berpengaruh terhadap kerawanan TBC

C. Analisis Spasial

- a. Integrasi Hasil Model dengan Data Geospasial. Pada tahap ini, output model disatukan dengan shapefile wilayah.
- b. Visualisasi Peta Risiko TBC menggunakan QGIS.
- c. Uji Autokorelasi Spasial:
 - Menggunakan Moran's I untuk mengukur pola distribusi spasial [16].
 - Getis-Ord Gi* digunakan untuk mengidentifikasi hot spot (zona merah) dan cold spot (zona aman) [17].

5. Alur Metode Penelitian

Adapun alur penelitian ini adalah sebagai berikut:



Gambar 1. Flowchart Penelitian

6. Perangkat Lunak dan Tools

- Python (Scikit-Learn, Pandas, Seaborn) digunakan untuk pemodelan machine learning.
- QGIS digunakan untuk pemetaan dan analisis spasial.
- Excel digunakan untuk pengolahan awal dan statistik deskriptif tambahan.

3. HASIL DAN PEMBAHASAN

A. Hasil Pemodelan Machine Learning

Penelitian ini menggunakan algoritma Random Forest Classifier untuk memprediksi wilayah rawan Tuberkulosis (TBC) berdasarkan beberapa variabel prediktor. Model dilatih dengan data dari 33 kabupaten/kota di Provinsi Sumatera Utara, dan hasil evaluasi model menunjukkan performa yang sangat baik.

- **Evaluasi Model**

Setelah proses pelatihan dan validasi menggunakan 5-fold cross-validation, diperoleh nilai metrik evaluasi sebagai berikut:

TABLE 1. Metrik Evaluasi

Metrik Evaluasi	Nilai
Akurasi	90,2%
Precision	86,7%
Recall	84,5%
F1-Score	84,1%

Nilai akurasi yang cukup tinggi menunjukkan bahwa model mampu mengklasifikasikan wilayah rawan dan tidak rawan dengan tingkat kesalahan yang rendah. F1-score yang mendekati nilai 1 menunjukkan keseimbangan antara precision dan recall, penting untuk klasifikasi yang sensitif terhadap kesalahan prediksi positif.

- **Feature Importance**

Berdasarkan analisis feature importance dari algoritma Random Forest, ditemukan bahwa variabel paling berpengaruh terhadap prediksi wilayah rawan TBC adalah:

1. Kepadatan Penduduk (28,3%)
2. Kualitas Hunian (23,5%)
3. Jumlah Fasilitas Kesehatan per Wilayah (20,1%)
4. Tingkat Kemiskinan (12,8%)
5. Status Gizi (9,7%)
6. Cakupan Imunisasi Dasar Lengkap (5,6%)

Variabel kepadatan penduduk menempati posisi teratas sebagai penentu risiko TBC. Hal ini sesuai dengan karakteristik penularan TBC yang tinggi pada lingkungan padat dan ventilasi buruk. Hunian yang tidak layak juga memperparah kondisi

lingkungan, serta fasilitas kesehatan yang terbatas menghambat deteksi dan pengobatan dini.

B. Hasil Analisis Spasial

Untuk memahami distribusi spasial dari hasil prediksi, dilakukan integrasi data prediktif dengan data spasial wilayah administrasi kabupaten/kota dan divisualisasikan dalam bentuk peta menggunakan QGIS.

• Pemetaan Risiko Wilayah

Peta klasifikasi wilayah dibagi dalam tiga kategori berdasarkan probabilitas prediksi:

1. Zona Merah (Risiko Sangat Tinggi)

- Kota Medan,
- Kabupaten Deli Serdang,
- Kabupaten Labuhanbatu

2. Zona Oranye (Risiko Tinggi)

- Kota Binjai
- Kabupaten Labuhan Batu Selatan
- Kota Padang Sidempuan
- Kabupaten Tapanuli Selatan
- Kabupaten Tapanuli Tengah
- Kabupaten Mandailing Natal
- Kabupaten Padang Lawas Utara
- Kabupaten Nias Selatan
- Kota Gunung Sitoli
- Kabupaten Serdang Bedagai
- Kabupaten Batu Bara
- Kota Tanjungbalai
- Kota Tebing Tinggi
- Kota Sibolga.

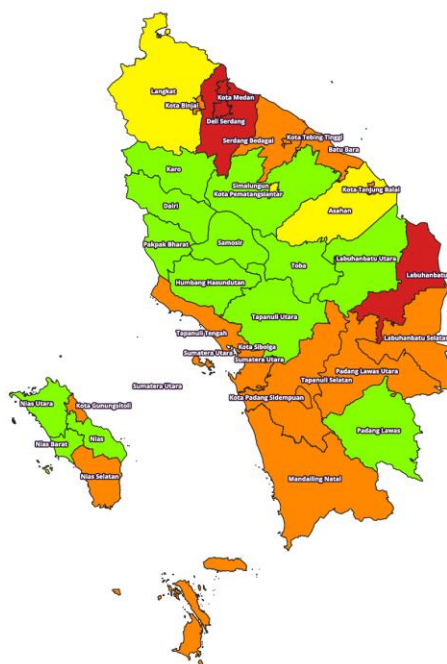
3. Zona Kuning (Risiko Sedang)

- Kota Pematangsiantar
- Kabupaten Langkat
- Kabupaten Asahan

4. Zona Hijau (Risiko Rendah)

- Kabupaten Nias
- Kabupaten Nias Utara
- Kabupaten Nias Barat
- Kabupaten Humbang Hasundutan
- Kabupaten Pakpak Bharat
- Kabupaten Labuhan Batu Utara
- Kabupaten Tapanuli Utara
- Kabupaten Padang Lawas
- Kabupaten Simalungun
- Kabupaten Dairi
- Kabupaten Humbang Hasundutan
- Kabupaten Toba
- Kabupaten Samosir
- Kabupaten Pakpak Bharat
- Kabupaten Karo

Peta ini sangat penting untuk mendukung pengambilan keputusan berbasis lokasi oleh Dinas Kesehatan, terutama dalam menentukan prioritas wilayah untuk intervensi aktif seperti skrining masif, edukasi, dan peningkatan akses layanan.



Gambar 2. Peta Spasial Kasus TBC Sumut

- **Autokorelasi Spasial**

Untuk memastikan apakah wilayah dengan risiko tinggi TBC tersebar secara acak atau terkonsentrasi, dilakukan uji autokorelasi spasial:

- **Moran's I = 0,47** ($p < 0,01$): menunjukkan **pola klaster positif**, artinya wilayah dengan risiko tinggi cenderung berdekatan secara geografis.
- **Getis-Ord Gi*** menghasilkan nilai signifikan ($p < 0,05$) untuk:
 - Kota Medan
 - Kabupaten Deli Serdang
 - Kabupaten Labuhanbatu

Hal ini menunjukkan adanya **hotspot spasial** TBC di wilayah tersebut. Temuan ini mendukung fakta bahwa kawasan perkotaan dan pinggiran dengan urbanisasi tinggi menjadi area konsentrasi kasus TBC.

C. Pembahasan

Hasil penelitian menunjukkan bahwa pendekatan integratif antara machine learning dan analisis spasial terbukti sangat efektif dalam mengidentifikasi wilayah berisiko tinggi terhadap tuberkulosis (TBC) di Provinsi Sumatera Utara. Model prediksi yang dikembangkan menggunakan algoritma Random Forest berhasil mencapai akurasi sebesar 90,2%, dengan hasil yang konsisten terhadap pola penyebaran TBC berdasarkan data historis dan spasial. Keunggulan pendekatan ini terletak pada kemampuannya tidak hanya menghasilkan prediksi yang kuat secara statistik, tetapi juga mengonversi hasil tersebut ke dalam bentuk visual spasial yang kontekstual dan mudah dipahami oleh pengambil kebijakan.

Faktor-faktor seperti kepadatan penduduk, kondisi tempat tinggal (overcrowding dan kualitas sanitasi), serta akses terhadap fasilitas kesehatan terbukti memiliki kontribusi besar dalam menentukan tingkat risiko suatu wilayah. Temuan ini sejalan dengan laporan WHO (2023) dan Kementerian Kesehatan RI (2022) yang menekankan pentingnya faktor sosial dan lingkungan dalam dinamika epidemi TBC. Selain itu, tingkat kemiskinan dan cakupan imunisasi BCG juga memiliki korelasi signifikan dalam mempengaruhi risiko, menunjukkan bahwa kerentanan sosial ekonomi memiliki peran yang tak kalah penting.

Peta risiko yang dihasilkan dalam penelitian ini mampu mengidentifikasi zona merah seperti Kota Medan, Kabupaten Deli Serdang, dan Kabupaten Labuhanbatu, yang selama ini memang dikenal memiliki beban kasus TBC yang tinggi. Keberadaan zona kuning dan zona hijau turut memperkuat urgensi pendekatan bertingkat (tiered intervention), di mana wilayah prioritas tinggi bisa mendapatkan sumber daya intervensi lebih besar dibandingkan wilayah dengan risiko rendah. Hal ini membuka peluang implementasi pendekatan precision public health, yakni strategi kesehatan masyarakat yang berbasis data dan menyasar intervensi spesifik sesuai karakteristik wilayah.

Lebih jauh, hasil analisis autokorelasi spasial dengan nilai Moran's I sebesar 0,47 ($p < 0,01$) menunjukkan adanya pola klastering geografis, yang berarti wilayah dengan risiko tinggi cenderung saling berdekatan. Temuan ini diperkuat dengan analisis Getis-Ord G_i^* yang berhasil mengidentifikasi hotspot signifikan di wilayah-wilayah kunci. Fakta ini penting karena menunjukkan bahwa intervensi di satu wilayah saja tidak cukup—dibutuhkan pendekatan regional yang memperhatikan konektivitas spasial antarwilayah.

Dengan demikian, penelitian ini memberikan dasar ilmiah yang kuat bagi penyusunan strategi pengendalian TBC berbasis wilayah dan data. Pemerintah daerah dapat menggunakan peta risiko sebagai alat bantu pengambilan keputusan, misalnya untuk:

- Menentukan lokasi prioritas skrining aktif dan pengobatan,
- Mengarahkan program rehabilitasi permukiman padat dan kumuh,
- Menyusun kampanye promosi kesehatan dan peningkatan kesadaran TBC secara lebih terfokus.

Integrasi metode machine learning dan spasial ini juga dapat direplikasi untuk penyakit menular lainnya dengan karakteristik geografis, sehingga berkontribusi besar terhadap penguatan sistem surveilans kesehatan digital di masa depan.

4. SIMPULAN

Penelitian ini berhasil mengintegrasikan algoritma machine learning, khususnya Random Forest, dengan analisis spasial untuk memprediksi dan memetakan wilayah rawan Tuberkulosis (TBC) di Provinsi Sumatera Utara. Hasil pemodelan menunjukkan bahwa model memiliki kinerja yang baik dengan akurasi sebesar 90,2%, serta mampu

mengidentifikasi variabel-variabel paling berpengaruh terhadap kerawanan TBC, yaitu kepadatan penduduk, kualitas hunian, dan akses terhadap fasilitas kesehatan.

Analisis spasial yang dilakukan melalui pemetaan risiko dan uji autokorelasi spasial memperkuat hasil model, dengan mengidentifikasi beberapa wilayah seperti Kota Medan, Kabupaten Deli Serdang, dan Kabupaten Labuhanbatu sebagai zona merah yang menunjukkan konsentrasi kasus TBC yang tinggi. Temuan ini mengindikasikan adanya pola kluster spasial yang signifikan.

Secara keseluruhan, pendekatan integratif ini mampu memberikan informasi yang lebih komprehensif, prediktif, dan berbasis lokasi, sehingga dapat digunakan sebagai dasar pengambilan keputusan yang lebih tepat sasaran dalam perencanaan dan implementasi program pengendalian TBC di Sumatera Utara. Penelitian ini juga menunjukkan bahwa pemanfaatan teknologi data sains memiliki potensi besar dalam mendukung kebijakan kesehatan masyarakat yang berbasis bukti (evidence-based policy).

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Negeri Medan (UNIMED) atas dukungan pendanaan penelitian yang diberikan. Dukungan ini sangat membantu dan berperan penting dalam kelancaran serta penyelesaian penelitian ini dengan baik.

REFERENCES

- [1] "View of Anti-TB Drug Side-Effects on the Treatment of Drug-Resistant Tuberculosis (DR-TB) in dr. Zainoel Abidin Hospital Banda Aceh." Accessed: Jun. 15, 2025. [Online]. Available: <https://jurnalrespirologi.org/index.php/jri/article/view/507/666>
- [2] A. N. Ayu Merzistya, M. Sakundarno Adi, D. Sutiningsih, and S. R. Rahayu, "THE QUALITY OF TUBERCULOSIS SERVICES IN PATIENTS' PERSPECTIVES: A LITERATURE REVIEW," *Indonesian Journal of Health Administration (Jurnal Administrasi Kesehatan Indonesia)*, vol. 9, no. 1, pp. 67–81, Jun. 2021, doi: 10.20473/JAKI.V9I1.2021.67-81.
- [3] "Global Tuberculosis Report 2023." Accessed: Jun. 15, 2025. [Online]. Available: <https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2023>
- [4] D. A. A. Zahra and B. Y. M. Sinaga, "Factors Affecting the Quality of Life Among Pulmonary Tuberculosis Patients at the TB-DOTS Outpatient Clinic of Public Hospital," *Public Health of Indonesia*, vol. 10, no. 1, pp. 43–51, Mar. 2024, doi: 10.36685/PHI.V10I1.774.
- [5] B. Y. M. Sinaga, D. A. Az Zahra, M. M. Simatupang, and A. S. Wahyuni, "Factors Associated with Quality of Life among Pulmonary Tuberculosis Patients," *Univers J Public Health*, vol. 12, no. 2, pp. 315–325, Apr. 2024, doi: 10.13189/UJPH.2024.120216.
- [6] I. Castiglioni *et al.*, "AI applications to medical images: From machine learning to deep learning," *Physica Medica: European Journal of Medical Physics*, vol. 83, pp. 9–24, Mar. 2021, doi: 10.1016/J.EJMP.2021.02.006.

-
- [7] F. Dodo Telaumbanua, P. Hulu, T. Zulfiter Nadeak, R. Romeo Lumbantong, and A. Dharma, "Penggunaan Machine Learning Di Bidang Kesehatan," *JURNAL TEKNOLOGI DAN ILMU KOMPUTER PRIMA (JUTIKOMP)*, vol. 2, no. 2, pp. 391–399, Oct. 2019.
 - [8] A. Satria, O. S. Sitompul, and H. Mawengkang, "5-Fold Cross Validation on Supporting K-Nearest Neighbour Accuration of Making Consimilar Symptoms Disease Classification," *Proceedings - 2nd International Conference on Computer Science and Engineering: The Effects of the Digital World After Pandemic (EDWAP), IC2SE 2021*, 2021, doi: 10.1109/IC2SE52832.2021.9792094.
 - [9] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," *Stata Journal*, vol. 20, no. 1, pp. 3–29, Mar. 2020, doi: 10.1177/1536867X20909688/ASSET/EC601327-319C-4C16-97CA-BC08F874F55A/ASSETS/IMAGES/10.1177_1536867X20909688-IMG9.PNG.
 - [10] M. Khairunisa, D. Made, S. A. Putri, G. Ngurah, and L. Wijayakusuma, "Perbandingan Metode Machine Learning untuk Analisis dan Prediksi Siklus Menstruasi," *JIEET (Journal of Information Engineering and Educational Technology)*, vol. 8, no. 2, pp. 111–115, May 2024, doi: 10.26740/JIEET.V8N2.P111-115.
 - [11] D. I. P. Desy, A. F. R. Kholdani, T. W. Qur'ana, and A. Dharmawati, "Pemodelan Spasial untuk Analisa Produksi Padi Integrasi Machine Learning," *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, vol. 14, no. 2, pp. 128–137, Nov. 2023, doi: 10.31849/DIGITALZONE.V14I2.16256.
 - [12] B. F. Khashoggi and A. Murad, "Issues of Healthcare Planning and GIS: A Review," *ISPRS International Journal of Geo-Information 2020, Vol. 9, Page 352*, vol. 9, no. 6, p. 352, May 2020, doi: 10.3390/IJGI9060352.
 - [13] F. Ramadhani, P. M. Fadillah, D. Septiana, A. Satria, and S. N. Amalia, "A Spatial Approach to Analyze the Distribution and Risk Factors of Stunting in North Sumatra with the K-Means Algorithm," *Proceedings of the International Conference on Electrical Engineering and Informatics*, pp. 1–6, 2024, doi: 10.1109/ICELTICS62730.2024.10776521.
 - [14] I. P. Sari, Al-Khowarizmi, F. Ramadhani, A. Satria, and O. K. Sulaiman, "Leukocoria Identification: A 5-Fold Cross Validation CNN and Adaboost Hybrid Approach," *6th International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2023 - Proceeding*, pp. 486–491, 2023, doi: 10.1109/ISRITI60336.2023.10467242.
 - [15] F. Ramadhani, M. Zarlis, and S. Suwilo, "Improve BIRCH algorithm for big data clustering," in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jan. 2020. doi: 10.1088/1757-899X/725/1/012090.
 - [16] A. Septiami *et al.*, "ANALISIS PENYEBARAN GIZI BURUK PADA BALITA DI NUSA TENGGARA BARAT (NTB) MENGGUNAKAN INDEKS MORAN (MORAN'S I) DAN LOCAL INDICATOR OF SPATIAL ASSOCIATION (LISA)," *Jurnal Gaussian*, vol. 13, no. 2, pp. 462–471, Dec. 2024, doi: 10.14710/J.GAUSS.13.2.462-471.
 - [17] M. Syukri *et al.*, "Identifikasi Pola Sebaran Spasial Dengan Metode Getis Ord Gi* Pada Kasus Tuberkulosis Di Kota Jambi 2018-2021," *Jurnal Ilmiah Permas: Jurnal Ilmiah STIKES Kendal*, vol. 12, no. 4, pp. 621–628, Oct. 2022, doi: 10.32583/PSKM.V12I4.712.