

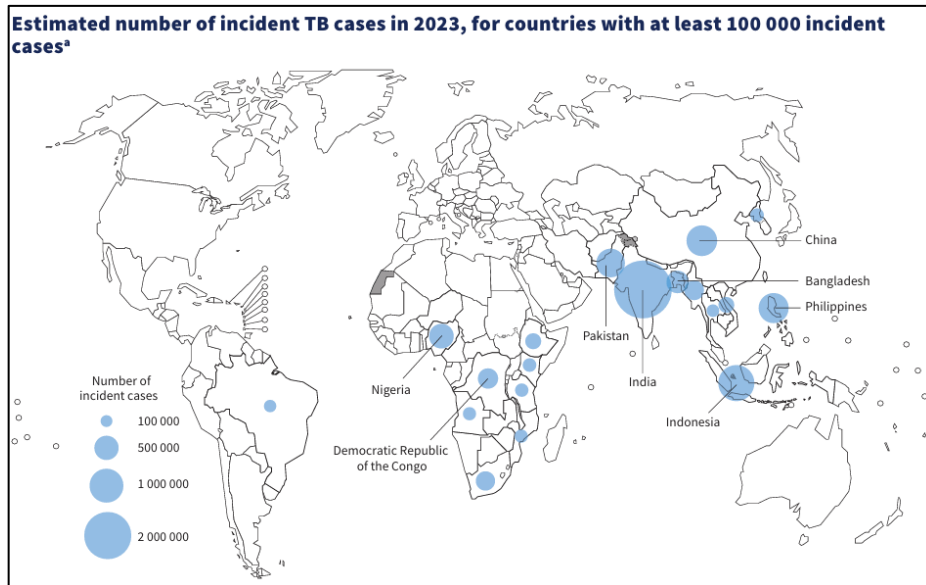
BAB I PENDAHULUAN

I.1 Latar Belakang

Paru-paru merupakan organ vital dalam sistem pernapasan manusia yang berfungsi untuk menyalurkan oksigen ke dalam darah dan mengeluarkan karbon dioksida dari tubuh. Organ ini memiliki jaringan alveolar yang luas, yang memungkinkan pertukaran gas secara efisien dan menjaga keseimbangan homeostatis tubuh (Laitupa & Amin, 2016). Gangguan terhadap fungsi paru-paru, dapat menimbulkan disfungsi respiratori yang berdampak sistemik pada kesehatan tubuh. Salah satu penyakit infeksi yang menyerang paru-paru dan menjadi masalah kesehatan global adalah tuberkulosis paru (TB paru), yang merupakan bentuk TB paling umum dan berperan besar dalam rantai penularan (Diantara dkk., 2022; Kementerian Kesehatan RI, 2025).

TB paru merupakan penyakit menular kronis yang disebabkan oleh *Mycobacterium tuberculosis* dan menyebar melalui udara saat penderita batuk atau mengeluarkan droplet infeksi (Inayah & Wahyono, 2019; Suhandi dkk., 2025). Risiko penularan TB paru meingkat pada individu dengan imunitas yang lemah, kondisi malnutrisi, kepadatan hunian tinggi, serta status sosial ekonomi rendah (Kementerian Kesehatan RI, 2025). Gejala yang timbul biasanya bersifat progresif, seperti batuk berkepanjangan (lebih dari dua hingga tiga minggu), hemoptisis, demam, keringat malam, penurunan berat badan yang tidak dapat dijelaskan, serta kelelahan kronis yang secara keseluruhan menurunkan kualitas hidup pasien (Ikatan Dokter Paru Indonesia, 2021).

Hal tersebut semakin relevan dengan data epidemiologi yang menunjukkan tingginya jumlah penderita TB paru di tingkat global maupun nasional. Berdasarkan laporan Organisasi Kesehatan Dunia (2024) dalam Yu dkk. (2025), pada tahun 2023 tercatat sekitar 10,8 juta kasus TB di seluruh dunia, dengan 1,25 juta kematian akibat penyakit ini. Selain itu, Indonesia juga menempati peringkat kedua jumlah kasus TB tertinggi di dunia setelah India, dengan estimasi 1 juta kasus atau sekitar 10% dari total global. Distribusi negara-negara dengan jumlah kasus TB terbanyak, termasuk Indonesia, dapat dilihat pada Gambar I-1.



Gambar I-1. Kasus TB secara global tahun 2023

Sumber: (World Health Organization, 2024)

Data nasional juga menunjukkan upaya yang semakin serius dalam pengendalian tuberkulosis di Indonesia. Berdasarkan rilis Kementerian Kesehatan RI tahun 2022, 91% kasus TB di Indonesia ialah TB paru yang berpotensi menular kepada orang terdekatnya dengan temuan sekitar 93.000 kasus kematian di tahun 2021 dan hanya sekitar 49% kasus yang berhasil ditemukan dan diobati (Kementerian Kesehatan RI, 2022). Pada tahun 2023, meskipun terjadi peningkatan estimasi kasus, sebanyak 821.200 kasus berhasil ditangani. Fakta ini menegaskan bahwa TB tetap menjadi beban kesehatan masyarakat yang serius dengan kesenjangan deteksi karena prevalensi TB di Indonesia yang cukup tinggi dari tahun ke tahun.

Selain tingginya jumlah kasus TB secara umum, perhatian khusus juga perlu diberikan terhadap komorbiditas yang menyertai TB paru. Komorbiditas yang sering ditemukan dalam TB paru ialah diabetes mellitus, HIV/AIDS, dan penyakit paru obstruktif kronik (PPOK) (Ikatan Dokter Paru Indonesia, 2021). Adanya komorbiditas dapat memperberat perjalanan klinis pasien, meningkatkan kegagalan pengobatan, memperpanjang masa penyembuhan, hingga meningkatkan angka kekambuhan dan mortalitas (Wijaya, 2015). Misalnya diabetes mellitus, yang menurunkan fungsi sistem imun tubuh sehingga meningkatkan kerentanan terhadap infeksi TB dan memperlambat penyembuhan jaringan paru. Penyakit lainnya, HIV/AIDS menyebabkan immunosupresi berat yang mempercepat aktivasi

TB laten menjadi TB paru aktif dengan gejala yang tidak khas. Sedangkan PPOK merusak struktur anatomi paru-paru dan mempermudah infeksi sekunder seperti TB paru. Jika penyakit komorbiditas yang menyertai TB paru ini terlambat didiagnosis, dapat menyebabkan komplikasi lebih lanjut. Sehingga, diperlukan identifikasi dini terhadap keberadaan komorbiditas untuk meningkatkan prognosis pasien.

Meskipun berbagai inisiatif kesehatan masyarakat telah dilakukan, identifikasi dini terhadap individu dengan risiko tinggi TB paru dengan komplikasi (komorbiditas) masih belum optimal. Metode diagnostik konvensional masih digunakan secara luas, namun sering kali kurang sensitif dan memerlukan waktu yang lama, sehingga berisiko menyebabkan salah diagnosis, keterlambatan pengobatan, dan peningkatan penularan penyakit (Martin dkk., 2021).

Sejalan dengan rencana pemerintah untuk meningkatkan efektivitas penanganan TB, salah satu pendekatan inovatif yang berkembang ialah pemanfaatan *machine learning*, di mana algoritma dapat mempelajari pola dari data klinis untuk melakukan prediksi penyakit. Model prediksi berbasis *machine learning* memiliki potensi besar sebagai *early warning system* bagi masyarakat untuk mengenali gejala dan risiko TB paru lebih awal, sekaligus alat bantu diagnosis bagi tenaga medis, khususnya dokter dalam mendukung pengambilan keputusan terapi.

Berbagai algoritma *machine learning* telah diterapkan dalam penelitian prediksi penyakit paru-paru. Penelitian yang dilakukan oleh Christian & Yani (2025) menggunakan algoritma Random Forest untuk memprediksi penyakit paru-paru dengan *dataset* berjumlah 30.000 data pasien. Penelitian tersebut menghasilkan nilai AUC sebesar 0,993, yang berarti tingkat akurasi sangat tinggi dalam klasifikasi penyakit paru-paru. Meskipun Random Forest mampu mengatasi *overfitting* dan menangani data berdimensi tinggi, algoritma ini memiliki keterbatasan dalam hal efisiensi komputasi jika berhadapan dengan jumlah data yang besar dan transparansi masing-masing atribut masih kurang dibanding algoritma *boosting* lainnya.

Pada penelitian lain oleh (Suhandi dkk., 2025) yang mengembangkan sistem klasifikasi penyakit tuberkulosis (TBC) menggunakan metode Uniform Manifold Approximation and Projection (UMAP) dan K-Nearest Neighbors (K-NN).

Penelitian ini menjelaskan bahwa penerapan UMAP sebagai metode reduksi dimensi berhasil meningkatkan akurasi klasifikasi K-NN dari 94,57% menjadi 100% dengan *precision*, *recall*, dan *F1-score* yang juga mencapai nilai sempurna. Namun, K-NN memiliki kelemahan inheren dalam mengolah *dataset* berdimensi tinggi tanpa reduksi dimensi sebelumnya. Selain itu, K-NN lebih sensitif terhadap pemilihan nilai K optimal dan distribusi data sehingga membuat model ini kurang stabil dibandingkan metode *ensemble modern*, khususnya pada data klinis yang lebih kompleks.

Sehubungan dengan metode *ensemble* yang dianggap lebih stabil dari penelitian sebelumnya, terbukti dengan penelitian yang dilakukan oleh Putra dkk. (2024) membandingkan efektivitas algoritma Random Forest, Extreme Gradient Boosting (XGBoost), dan Logistic Regression dalam klasifikasi penyakit paru-paru. Hasil penelitian menunjukkan performa XGBoost lebih unggul dibandingkan dengan dua algoritma lainnya, yakni akurasi sebesar 94,44%, *precision* 94,98%, *recall* 94,44% dan *F1-score* 94,41% setelah dilakukannya *hyperparameter tuning*. Sedangkan Random Forest memperoleh akurasi 94,33% dan Logistic Regression sebesar 87,40%. Meskipun Random Forest dan XGBoost efektif dalam klasifikasi penyakit paru-paru, XGBoost lebih unggul dalam menyeimbangkan *precision* dan *recall*. Hal ini mempertegas keunggulan algoritma berbasis pohon keputusan dalam klasifikasi penyakit berbasis data tabular yang kompleks.

Efektivitas XGBoost dalam diagnosis penyakit telah dibuktikan melalui penelitian lain. Dalam penelitian oleh Shu dkk. (2024) menunjukkan bahwa model XGBoost mampu membedakan penyakit Crohn dan tuberkulosis intestinal dengan nilai ROC-AUC sebesar 0,946 dan akurasi 88,4%. Model ini juga diuji dalam praktik klinis nyata dengan metode interpretabilitas SHAP dan LIME yang mencapai tingkat kesesuaian 90,7%. Selain itu, (Ghosh & Khandoker, 2024) juga membuktikan bahwa XGBoost memiliki performa tertinggi dalam prediksi penyakit ginjal kronis dibanding lima algoritma lainnya, dengan AUC 0,9689 dan akurasi 93,29% yang menekankan pentingnya interpretasi klinis menggunakan SHAP dan LIME dalam mendukung keputusan medis. Temuan ini memperkuat pemilihan XGBoost sebagai algoritma yang tidak hanya akurat, tetapi dapat menginterpretasikan dengan baik.

Berbeda dengan penelitian-penelitian sebelumnya yang sebagian besar hanya berfokus pada klasifikasi TB paru secara umum atau penyakit paru-paru sebagai data utama, penelitian ini menekankan pada prediksi penyakit TB paru dan TB paru dengan komorbiditas berdasarkan faktor klinis seperti data demografis, gejala klinis (seperti batuk berdahak, demam, sesak napas, dan lainnya), hasil laboratorium (leukosit, trombosit, LED, tes BTA, tes TCM, dan lainnya), dan faktor risiko pasien (riwayat kontak TB, riwayat merokok, dan lainnya) yang mencerminkan kondisi fisiologis pasien dari rekam medis pasien RSUD Al-Ihsan Bandung.

Pemilihan algoritma XGBoost didasarkan pada keunggulannya dalam mengolah data tabular dengan atribut numerik dan kategorikal, serta mampu menghasilkan model yang interpretatif melalui *feature importance*. Dengan hal tersebut, penelitian ini juga mengusulkan pengembangan *dashboard* interaktif yang dapat mendukung intervensi medis dan pengelolaan risiko pasien TB paru dengan menyajikan hasil prediksi secara visual dan mudah dipahami. Selain itu, *dashboard* interaktif menjadi sistem peringatan dini bagi masyarakat untuk mengenali faktor risiko penyakit TB paru, khususnya dengan penyakit komorbiditas agar dapat melakukan pencegahan atau pemeriksaan secara berkala jika mengalami gejala penyakit tersebut. Sehingga dengan pendekatan *machine learning*, penelitian ini berupaya memberikan kontribusi nyata dalam pengelolaan dan pencegahan penyakit secara menyeluruh.

I.2 Rumusan Masalah

Rumusan masalah dari permasalahan yang terdapat pada latar belakang adalah sebagai berikut:

1. Bagaimana hasil model klasifikasi penyakit TB paru dan TB paru dengan komorbiditas berdasarkan data rekam medis pasien menggunakan algoritma XGBoost?
2. Berdasarkan hasil *feature importance* dari algoritma XGBoost, apa saja faktor klinis yang memengaruhi klasifikasi penyakit TB paru dan TB paru dengan komorbiditas?

3. Bagaimana sebuah *dashboard* interaktif dapat memvisualisasikan hasil pemodelan klasifikasi dan analisis faktor klinis dari penyakit TB paru dan TB paru dengan komorbiditas?

I.3 Tujuan Tugas Akhir

Tujuan penelitian dari permasalahan yang telah dijelaskan pada rumusan masalah ini adalah sebagai berikut:

1. Membangun dan menguji model klasifikasi pasien TB paru dan TB paru dengan komorbiditas menggunakan XGBoost berbasis data rekam medis pasien yang diperoleh dari RSUD Al-Ihsan Bandung.
2. Menganalisis faktor klinis yang memengaruhi klasifikasi penyakit TB paru dan TB paru dengan komorbiditas menggunakan *feature importance* dari metode SHAP.
3. Mengembangkan *dashboard* interaktif berbasis *web* menggunakan Streamlit untuk memprediksi penyakit dan memvisualisasikan hasil klasifikasi dan faktor klinis risiko penyakit TB paru dengan komorbiditas.

I.4 Manfaat Tugas Akhir

Berdasarkan tujuan penelitian yang akan dijelaskan, manfaat yang diharapkan dari penelitian ini sebagai berikut:

1. Bagi Tenaga Medis

Hasil model klasifikasi dan analisis faktor klinis dapat dimanfaatkan sebagai alat bantu dalam proses diagnosis, sehingga membantu tenaga medis mengenali kondisi pasien secara lebih efisien dan efektif. Hal tersebut mendukung pengambilan keputusan klinis yang lebih tepat dalam menentukan prioritas penanganan pasien.

2. Bagi Institusi Layanan Kesehatan

Dashboard interaktif yang dikembangkan dimanfaatkan sebagai sistem pendukung keputusan yang menyajikan visualisasi data dengan informatif. Institusi dapat meningkatkan efisiensi dalam proses diagnosis dan pengelolaan sumber daya medis dengan tepat sasaran. Selain itu, penerapan teknologi juga

menjadi langkah awal menuju transformasi digital dalam layanan kesehatan, khususnya mencegah penyakit TB paru dengan komorbiditas.

3. Bagi Masyarakat

Berkontribusi secara tidak langsung dalam meningkatkan kesadaran masyarakat akan pentingnya pemeriksaan rutin dan skrining awal penyakit TB paru, terutama pada faktor risiko komorbiditas. Selain itu, dapat memperoleh akses informasi kesehatan yang efisien melalui faktor klinis dan mendorong inisiatif pemeriksaan lebih lanjut melalui layanan kesehatan.

I.5 Batasan dan Asumsi Tugas Akhir

Batas penelitian pada permasalahan ini adalah sebagai berikut:

1. *Dataset* yang digunakan berdasarkan data rekam medis pasien rawat jalan pada RSUD Al-Ihsan Bandung tahun 2024.
2. Parameter yang akan digunakan terbatas pada usia, jenis kelamin, gejala klinis, Laju Endapan Darah (LED), *C-Reactive Protein* (CRP), leukosit, hemoglobin, trombosit, dan parameter lainnya.
3. Mengklasifikasikan penyakit TB paru dengan dan tanpa komorbiditas (penyakit penyerta) dengan evaluasi performa model klasifikasi menggunakan *confusion matrix* dan *log-loss*.
4. *Dashboard* interaktif terbatas pada hasil prediksi penyakit yang disertai dengan probabilitas diagnosis penyakit dan saran penanganan medis serta visualisasi faktor klinis risiko penyakit TB paru dengan komorbiditas.

I.6 Sistematika Laporan

Penelitian ini diuraikan dengan sistematika penulisan sebagai berikut:

Bab I Pendahuluan

Pada bab ini berisi uraian mengenai konteks permasalahan, latar belakang penelitian, perumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penulisan.

Bab II Landasan Teori

Bab ini berisi literatur yang relevan dengan permasalahan yang diteliti dan dibahas hasil-hasil penelitian terdahulu. Minimal terdapat lebih dari satu metodologi/metode/kerangka kerja yang disertakan pada bab ini untuk menyelesaikan permasalahan atau meminimalisir gap antara kondisi eksisting dengan target. Pada akhir bab, analisis pemilihan metodologi/metode/kerangka kerja harus dijelaskan untuk menentukan metodologi/metode/kerangka kerja yang akan digunakan di penelitian ini.

Bab III Metode Penyelesaian Masalah

Bab ini menjelaskan pendekatan, metode, dan prosedur yang digunakan untuk merumuskan serta mengimplementasikan solusi atas masalah yang telah diidentifikasi. Isi bab meliputi desain dan tahapan penelitian secara sistematis untuk mencapai tujuan tugas akhir, termasuk kerangka pikir, parameter, model yang digunakan, metode pengumpulan dan analisis data, serta langkah-langkah penelitian. Penjelasan bab ini dapat dilengkapi dengan diagram alir atau gambar pendukung, dan jika ada wawancara, daftar pertanyaan dilampirkan.

Bab IV Penyelesaian Permasalahan

Bab ini merupakan penjelasan mengenai proses pengumpulan dan pengolahan data yang diperlukan untuk menyelesaikan masalah rekayasa dengan mempertimbangkan sistem terintegrasi. Bab ini mencakup detail metode yang digunakan untuk mengumpulkan data, teknik pengolahan data dan bagaimana proses data tersebut diterapkan untuk dilakukan analisis dan pengambilan keputusan.

Bab V Validasi, Analisis Hasil, dan Implikasi

Bab ini berisikan uraian proses validasi serta analisis hasil penelitian yang dilakukan untuk menyelesaikan permasalahan. Selanjutnya, secara terperinci dan tahap demi tahap tujuan penelitian dibahas dan dianalisis secara detail dan tajam, dengan menggunakan metode yang telah diberikan dalam metodologi penelitian, hingga diperoleh suatu hasil penelitian.

Bab VI Kesimpulan dan Saran

Bagian terakhir dari penelitian tugas akhir ini yang bertujuan untuk membuat kesimpulan dari hasil penelitian dan memberikan saran sesuai. bagian ini mencerminkan kontribusi penelitian terhadap pemahaman dan penyelesaian masalah.