

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Saat ini, sektor kesehatan global menghadapi tantangan besar akibat meningkatnya beban penyakit infeksi yang disertai komplikasi berat. Salah satu kondisi yang memiliki tingkat mortalitas tinggi adalah sepsis, yaitu disfungsi organ yang mengancam jiwa akibat respons tubuh yang tidak teratur terhadap infeksi [1]. Berdasarkan data tahun 2021, diperkirakan terdapat sekitar 166 juta kasus sepsis di seluruh dunia dengan angka kematian mencapai 21,4 juta jiwa atau sekitar 31,5% dari seluruh kematian global [2]. Beban penyakit ini masih didominasi oleh negara-negara berpenghasilan rendah dan menengah (*Low-Middle Income Countries* - LMICs), termasuk kawasan Asia Tenggara dan Afrika Sub-Sahara [3]. Tingginya angka kejadian dan mortalitas tersebut menjadikan sepsis sebagai salah satu penyebab utama kematian pada pasien yang dirawat di unit perawatan intensif (*Intensive Care Unit* - ICU) serta menjadi tantangan penting dalam pelayanan kesehatan modern [4].

Di antara berbagai kelompok pasien, penderita kanker merupakan populasi yang memiliki risiko lebih tinggi mengalami sepsis dibandingkan populasi umum [5, 6]. Risiko tersebut dipengaruhi oleh kondisi immunosupresi akibat perkembangan penyakit kanker maupun efek terapi seperti kemoterapi, radioterapi, dan terapi target yang menyebabkan penurunan sistem kekebalan tubuh [4]. Kondisi tersebut meningkatkan kerentanan terhadap infeksi berat yang dapat berkembang menjadi sepsis. Diperkirakan sekitar 20% pasien yang dirawat di ICU memiliki riwayat penyakit onkologi [6]. Selain itu, analisis nasional menunjukkan lebih dari 684.930 kematian terkait sepsis pada pasien dewasa dengan kanker, dengan angka kematian yang disesuaikan berdasarkan usia (*Age-Adjusted Mortality Rate* - AAMR) meningkat dari 12,06 menjadi 14,21 dalam dua dekade terakhir [7]. Fakta tersebut menunjukkan bahwa pasien kanker dengan sepsis merupakan kelompok berisiko tinggi yang memerlukan penanganan dan identifikasi prognosis secara cepat serta akurat.

Prediksi mortalitas pada pasien kanker dengan sepsis menjadi aspek penting dalam mendukung pengambilan keputusan klinis, seperti penentuan tingkat intensitas perawatan, prioritas penggunaan sumber daya ICU, hingga penyusunan

strategi terapi yang lebih tepat. Salah satu indikator prognosis yang banyak digunakan dalam penelitian klinis adalah mortalitas dalam 28 hari (*28-day mortality*) karena mampu merepresentasikan luaran klinis jangka pendek setelah pasien mengalami episode sepsis. Identifikasi dini terhadap pasien yang berisiko mengalami kematian dalam periode tersebut berpotensi membantu tenaga medis melakukan intervensi lebih cepat sehingga dapat meningkatkan kualitas pelayanan serta menurunkan angka mortalitas.

Berbagai sistem penilaian klinis konvensional seperti Sequential Organ Failure Assessment (SOFA), Simplified Acute Physiology Score II (SAPS II), maupun Acute Physiology and Chronic Health Evaluation (APACHE) telah banyak digunakan untuk memperkirakan tingkat keparahan pasien kritis. Namun demikian, sistem penilaian tersebut memiliki keterbatasan dalam menangkap hubungan kompleks dan non-linear antar berbagai variabel klinis, laboratorium, serta karakteristik pasien, khususnya pada populasi pasien kanker dengan sepsis yang memiliki kondisi fisiologis lebih heterogen. Oleh karena itu, diperlukan pendekatan analitik yang mampu mengolah data klinis secara lebih komprehensif untuk meningkatkan akurasi prediksi mortalitas.

Perkembangan *Machine Learning* memberikan peluang baru dalam membangun model prediksi mortalitas berdasarkan data klinis pasien. Berbagai algoritma seperti Extreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Decision Tree, Random Forest, maupun Logistic Regression telah banyak diterapkan pada data kesehatan karena mampu mempelajari pola kompleks dari data multidimensional serta menghasilkan performa prediksi yang kompetitif [8, 9]. Pemanfaatan parameter laboratorium, tanda vital, karakteristik demografis, serta informasi klinis lainnya memungkinkan model *machine learning* mengidentifikasi faktor-faktor yang berkaitan dengan risiko kematian secara lebih efektif dibandingkan pendekatan statistik konvensional.

Meskipun berbagai penelitian telah mengembangkan model prediksi mortalitas pada pasien sepsis, penelitian yang secara khusus berfokus pada populasi pasien kanker dengan sepsis masih relatif terbatas. Selain itu, sebagian besar penelitian hanya mengevaluasi satu atau dua algoritma sehingga belum memberikan gambaran komprehensif mengenai performa berbagai metode *machine learning* pada populasi tersebut. Oleh karena itu, diperlukan penelitian yang membandingkan beberapa algoritma klasifikasi untuk menentukan model dengan performa terbaik dalam memprediksi mortalitas 28 hari pada pasien kanker dengan sepsis.

Penelitian ini bertujuan membangun dan mengevaluasi beberapa model *machine learning* untuk memprediksi mortalitas dalam 28 hari pada pasien kanker dengan sepsis menggunakan data klinis yang diperoleh dari basis data Medical Information Mart for Intensive Care (MIMIC-IV). Tahapan penelitian meliputi eksplorasi karakteristik data (*Exploratory Data Analysis*), proses prapemrosesan data, pembangunan model klasifikasi menggunakan beberapa algoritma *machine learning*, serta evaluasi performa model menggunakan metrik klasifikasi yang sesuai. Hasil penelitian diharapkan dapat memberikan rekomendasi algoritma dengan performa terbaik sebagai dasar pengembangan sistem pendukung keputusan klinis dalam mengidentifikasi pasien kanker dengan sepsis yang berisiko tinggi mengalami kematian dalam 28 hari.

1.2 Rumusan Masalah

1. Bagaimana performa beberapa algoritma machine learning dalam memprediksi mortalitas 28 hari pada pasien kanker dengan sepsis menggunakan data klinis dari MIMIC-IV?
2. Faktor klinis apa saja yang paling berkontribusi terhadap prediksi mortalitas 28-hari pada pasien kanker dengan sepsis berdasarkan hasil analisis model machine learning?

1.3 Batasan Permasalahan

Terdapat beberapa Batasan masalah pada penelitian yang dilakukan, yaitu:

1. Penelitian ini menggunakan data sekunder yang berasal dari database *Medical Information Mart for Intensive Care IV* (MIMIC-IV) versi 3.1 dan hanya mencakup pasien dewasa (usia ≥ 18 tahun) yang memenuhi kriteria Sepsis-3 serta memiliki kanker aktif selama perawatan di Intensive Care Unit (ICU).
2. Penelitian ini hanya menggunakan variabel klinis yang diperoleh dalam 24 jam pertama perawatan ICU. Variabel dengan persentase *missing value* lebih dari 20% dikeluarkan dari analisis, sedangkan *missing value* pada variabel yang memenuhi kriteria ditangani menggunakan metode *Multiple Imputation by Chained Equations* (MICE).
3. Penelitian ini berfokus pada pengembangan dan evaluasi model *machine learning* untuk memprediksi mortalitas 28 hari pada pasien kanker dengan

sepsis. Penelitian ini tidak bertujuan untuk menganalisis hubungan sebab-akibat maupun mengevaluasi efektivitas terapi atau intervensi klinis.

1.4 Tujuan Penelitian

Terdapat tujuan penelitian yang telah ditetapkan, yaitu:

1. Mengevaluasi dan membandingkan performa beberapa algoritma machine learning dalam memprediksi mortalitas 28-hari pada pasien kanker dengan sepsis menggunakan data MIMIC-IV.
2. Mengidentifikasi faktor-faktor klinis yang paling berpengaruh terhadap prediksi mortalitas 28-hari berdasarkan hasil analisis feature importance.

1.5 Manfaat Penelitian

Terdapat manfaat penelitian yang diharapkan, yaitu:

1. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan penerapan metode *machine learning* untuk prediksi mortalitas 28 hari pada pasien kanker dengan sepsis. Selain itu, penelitian ini dapat menjadi referensi bagi penelitian selanjutnya yang memanfaatkan data MIMIC-IV dalam pengembangan model prediksi klinis pada pasien kritis.
2. Penelitian ini diharapkan dapat menjadi dasar dalam pengembangan sistem pendukung keputusan klinis (*Clinical Decision Support System*) untuk membantu tenaga kesehatan mengidentifikasi pasien kanker dengan sepsis yang memiliki risiko mortalitas tinggi sehingga dapat mendukung penentuan prioritas pemantauan dan penatalaksanaan pasien di ICU.

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN

Bab ini menguraikan gambaran umum penelitian yang meliputi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan laporan.

- Bab 2 LANDASAN TEORI

Bab ini membahas tentang teori teori yang mendukung dan kajian pustaka.

- Bab 3 METODOLOGI PENELITIAN

Bab ini menjelaskan metode yang digunakan, *pre processing* dataset, pelatihan model, serta metode evaluasi model.

- Bab 4 HASIL DAN DISKUSI

Bab ini berisi hasil dan pembahasan penelitian berdasarkan evaluasi model yang dibuat.

- Bab 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang diperoleh dari hasil penelitian serta saran untuk penelitian selanjutnya.

