

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sistem pelayanan kesehatan saat ini menghadapi tantangan yang semakin besar terkait manajemen pasien dengan peningkatan kompleksitas kondisi klinis. Salah satu yang menjadi faktor penyebab kondisi tersebut ialah tingginya prevalensi multimorbiditas. Multimorbiditas ialah situasi dari seseorang dengan dua atau lebih kondisi medis secara bersamaan tanpa adanya satu kondisi medis yang menjadi fokus primer (France *et al.*, 2012). Multimorbiditas juga dapat diartikan sebagai koeksistensi dari dua atau lebih kondisi medis, baik penyakit kronis maupun akut pada individu tertentu yang terjadi di waktu bersamaan (Rumawas dan Buchori, 2023). Prevalensi multimorbiditas meningkat di berbagai populasi, hal ini didorong oleh penuaan, perubahan gaya hidup, peningkatan status ekonomi-sosial, dan peningkatan kemampuan diagnostik layanan kesehatan. Tingginya prevalensi multimorbiditas mengakibatkan konsekuensi negatif, seperti meningkatnya angka mortalitas serta meningkatnya penggunaan dan biaya layanan kesehatan (Chowdhury *et al.*, 2023).

Multimorbiditas telah berkembang menjadi isu kesehatan di masyarakat karena dampaknya yang luas terhadap kualitas hidup. Pentingnya penelitian mengenai kondisi ini tercermin dari temuan Chowdhury *et al.* (2023) yang menganalisis data dari 126 publikasi dengan melibatkan 15,4 juta orang (32,1% diantaranya adalah laki-laki) yang berusia rata-rata 56,94 tahun (simpangan baku 10,84 tahun) dari 54 negara di seluruh dunia. Hasil yang diperoleh dari temuan ini

adalah prevalensi multimorbiditas secara global mencapai 37,2%. Temuan ini selaras dengan hasil analisis Violan *et al.* (2014) yang mengidentifikasi 39 publikasi dan melibatkan 70.057.611 pasien di 12 negara, diperoleh prevalensi multimorbiditas berkisar 12,9% pada pasien berusia 18 tahun ke atas serta 95,1% pada pasien berusia 65 tahun ke atas.

Temuan di Indonesia oleh Mahwati (2014) menyatakan bahwa analisis yang dilakukan terhadap data sekunder survei *Indonesian Family Life Survey* (IFLS) keempat pada 2007 diperoleh hasil prevalensi multimorbiditas sebesar 15,8% dari total sampel 2.960 orang yang berusia ≥ 60 tahun. Temuan ini sesuai dengan laporan The National Institute of Health Research and Development (2008) yang menyebutkan bahwa hanya 13% lansia berada dalam kondisi sehat, sementara 34,8% menderita satu jenis penyakit dan 52,2% menderita dua jenis penyakit atau lebih (multimorbiditas). Keselarasan temuan ini menegaskan bahwa multimorbiditas merupakan permasalahan kesehatan yang penting untuk dilakukan penelitian.

Data rekam medis pasien seringkali menunjukkan adanya kondisi multimorbiditas. Fenomena ini memungkinkan terbentuknya berbagai kombinasi penyakit yang muncul secara simultan, tetapi penelitian ini menetapkan batasan analisis pada kelompok penyakit tertentu. Penelitian ini berfokus pada penyakit stroke, gagal jantung, penyakit jantung iskemik, gagal ginjal, infeksi saluran kemih, dan demam tifoid. Kombinasi yang terjadi dari penyakit tersebut dapat meningkatkan risiko komplikasi dan mortalitas pada pasien. Adanya kombinasi dari penyakit tersebut menuntut keputusan klinis yang cepat, akurat, dan berbasis data. Guna mengelola kompleksitas klinis secara efektif dan efisien, instrumen

pendukung keputusan membutuhkan kapabilitas klasifikasi untuk melakukan prediksi terhadap serangkaian penyakit. *Machine learning* akan membantu tenaga medis dalam melakukan klasifikasi terhadap serangkaian penyakit dengan lebih baik. *Machine learning* dapat mengurangi variabilitas keputusan tenaga medis dalam melakukan diagnosis penyakit sehingga menjadi lebih konsisten (Frénay dan Verleysen, 2014).

Klasifikasi merupakan tindakan menilai objek data untuk menempatkannya ke suatu kelas dari beberapa kelas yang ada (Lishania *et al.*, 2019). Penerapannya pada data rekam medis melibatkan penentuan kategori penyakit yang sesuai untuk setiap pasien berdasarkan diagnosis yang ada. Data rekam medis pasien multimorbiditas bersifat *multi-label*, dimana setiap pasien dapat mempunyai lebih dari satu penyakit secara bersamaan. Penggunaan teknik klasifikasi tradisional yang hanya dapat memproyeksikan satu hasil tunggal tidak dapat digunakan. Guna menangani permasalahan tersebut diperlukan perluasan teknik klasifikasi, yaitu penerapan klasifikasi *multi-label*.

Klasifikasi *multi-label* adalah metode klasifikasi untuk membangun model yang mampu mempelajari pola dari data dan melakukan prediksi kombinasi label berdasarkan fitur yang ada. Klasifikasi *multi-label* bisa diterapkan menggunakan pendekatan *Problem Transformation* dan *Algorithm Adaptation* (Zhang dan Zhou, 2014). Pendekatan *Problem Transformation* bekerja dengan mengubah permasalahan *multi-label* menjadi satu atau lebih permasalahan klasifikasi *single-label* sedangkan pendekatan *Algorithm Adaptation* memodifikasi algoritma klasifikasi yang sudah ada agar dapat menangani data *multi-label* secara langsung.

Penelitian ini menerapkan salah satu pendekatan *Problem Transformation*, yaitu *Binary Relevance*. Cara kerja dari *Binary Relevance* adalah dengan menerapkan konsep *binary classifier* di setiap label kemudian dilakukan penggabungan hasil untuk memperoleh himpunan label prediksi (Zhang *et al.*, 2018). *Binary Relevance* merupakan metode *Problem Transformation* yang sederhana, mudah untuk diterapkan, dan fleksibel karena mampu dipadukan dengan berbagai algoritma klasifikasi. Algoritma yang diterapkan dalam penelitian ini adalah *Naive Bayes*. Penerapan algoritma ini disebabkan karena *Naive Bayes* dapat menentukan parameter dengan baik dalam proses klasifikasi meskipun menggunakan data pelatihan yang relatif sedikit (Rifai *et al.*, 2019). Algoritma *Naive Bayes* menerapkan teorema bayes dan mengasumsikan bahwa semua fitur saling independen apabila kondisi label sudah diketahui (Syarli dan Muin, 2016).

Penelitian terdahulu oleh Rizkyani *et al.* (2022) mengenai analisis klasifikasi *multi-label* terkait penyakit kanker serviks menggunakan dataset *Cervical Cancer* yang terdiri dari data pasien ‘*Hospital Universitario de Caracas*’ di Caracas, Venezuela. Penelitian dilakukan dengan menerapkan algoritma ML-KNN. Hasil penelitian diperoleh bahwa penerapan ML-KNN dapat memberikan performa yang baik pada klasifikasi *multi-label* untuk penyakit kanker serviks menggunakan dataset *Cervical Cancer*. Penelitian lain yang dilakukan oleh Virginio dan dos Reis (2018) mengenai klasifikasi *multi-label* untuk melakukan pengodean ICD secara otomatis dari ringkasan pemulangan pasien yang ditulis dalam bentuk teks bebas. Dataset yang digunakan dalam penelitian ini adalah MIMIC-III yang berisi dari 53.018 ringkasan pemulangan pasien rawat inap dewasa. Penelitian ini menggunakan metode *Binary Relevance* dengan algoritma SVM. Di sisi lain,

penelitian ini juga mengeksplorasi peran dari optimasi *Binary Relevance* dengan algoritma SVM yang dikombinasikan dengan *class weighting* untuk menangani ketidakseimbangan label. Hasil dari penelitian ini adalah penggunaan *class weighting* dapat memberikan performa yang lebih baik dibandingkan sebelum menggunakan *class weighting*. Yazid *et al.* (2022) juga melakukan analisis klasifikasi *multi-label* mengenai deteksi ujaran kebencian dengan menggunakan dataset *multi-label twitter* berbahasa Indonesia dengan total label sebanyak empat. Metode yang diterapkan adalah *Label Powerset* yang digabungkan dengan algoritma *Naive Bayes*, fitur *n-gram*, dan pembobotan kata TF-IDF. Hasil penelitian ini menyatakan bahwa penerapan metode *Label Powerset* yang digabungkan dengan algoritma *Naive Bayes*, fitur *n-gram*, dan pembobotan kata TF-IDF dapat memberikan performa baik dalam deteksi ujian kebencian dengan dataset *multi-label twitter* berbahasa Indonesia.

Berdasarkan penelitian terdahulu terdapat beberapa metode yang diterapkan dalam analisis klasifikasi *multi-label*, seperti ML-KNN, *Binary Relevance* dengan algoritma SVM, dan *Label Powerset* dengan algoritma *Naive Bayes*. Penelitian yang membahas kasus multimorbiditas pada analisis klasifikasi masih terbatas, meskipun kondisi multimorbiditas merupakan masalah penting dalam bidang kesehatan dengan tingkat kompleksitas tinggi dan memerlukan kajian yang lebih mendalam. Penggunaan metode *Binary Relevance* dengan algoritma *Naive Bayes* terutama dalam analisis multimorbiditas juga masih jarang dilakukan. Berdasarkan kondisi tersebut, penelitian ini berfokus pada penerapan metode *Binary Relevance* dengan algoritma *Naive Bayes* menggunakan data rekam medis pasien berdasarkan

kondisi multimorbiditas untuk mengklasifikasikan jenis penyakit yang dialami serta melakukan evaluasi performa untuk mengetahui kinerja dari algoritma tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, permasalahan dalam penelitian ini dapat dirumuskan sebagai berikut.

1. Bagaimana penerapan algoritma *Naive Bayes* pada analisis klasifikasi *multi-label* menggunakan data pasien berdasarkan kondisi multimorbiditas?
2. Bagaimana performa model analisis klasifikasi *multi-label* pada pasien berdasarkan kondisi multimorbiditas menggunakan algoritma *Naive Bayes*?

1.3 Batasan Masalah

Batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut.

1. Pendekatan yang diterapkan dalam klasifikasi *multi-label* adalah *Problem Transformation* dengan metode *Binary Relevance*.
2. Algoritma klasifikasi yang digunakan adalah *Naive Bayes*.
3. Tidak mencakup tahap pengecekan serta penanganan *imbalanced* data.
4. Pembagian data latih dan data uji dilakukan dengan tiga jenis, yaitu menggunakan proporsi 80%:20%, 70%:30%, serta 60%:40%.

1.4 Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah yang telah dipaparkan, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Menerapkan algoritma *Naive Bayes* pada analisis klasifikasi *multi-label* menggunakan data pasien berdasarkan kondisi multimorbiditas.
2. Mengevaluasi performa model analisis klasifikasi *multi-label* pada pasien berdasarkan kondisi multimorbiditas menggunakan algoritma *Naive Bayes*.