

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang Masalah**

Kesehatan global saat ini sedang menghadapi krisis serius akibat meningkatnya jumlah kematian yang disebabkan oleh penyakit kronis yang terus bertambah setiap tahunnya. Perubahan gaya hidup, pola makan, dan faktor lingkungan telah berkontribusi pada meningkatnya risiko berbagai penyakit, baik penyakit yang bersifat menular maupun yang tidak menular [1]. Diabetes, gagal jantung, dan kanker termasuk dalam kelompok penyakit tidak menular yang kini menduduki peringkat teratas sebagai penyebab kematian di banyak negara. Pada tahun 2023, kelompok penyakit ini diperkirakan berkontribusi terhadap sekitar 74% kematian global, di mana diabetes, gagal jantung, dan kanker menjadi faktor penyumbang terbesar [2].

Berdasarkan data *World Health Organization (WHO)* tahun 2022 jumlah penderita diabetes di dunia mencapai 830 juta jiwa, dengan angka kematian akibat penyakit tersebut diperkirakan mencapai 3,4 juta jiwa secara global. Sementara itu, data *Global Cancer Observatory* tahun 2022 menunjukkan bahwa jumlah penderita kanker mencapai 20 juta jiwa, dengan estimasi kematian sebesar 9,7 juta jiwa. Adapun menurut *World Heart Federation* pada tahun 2022, jumlah penderita gagal jantung diperkirakan mencapai 64 juta jiwa, dan sekitar 50% dari mereka yang terdiagnosis gagal jantung meninggal dalam kurun waktu lima tahun setelah diagnosis.

Tingginya angka penderita dan kematian akibat penyakit tersebut menuntut adanya sistem pencatatan dan pengelolaan medis yang akurat, mengingat kesalahan dalam dokumentasi dapat menyebabkan keterlambatan diagnosis, pengobatan yang tidak tepat, hingga meningkatkan risiko kematian pasien [3]. Namun, kesalahan dalam pendokumentasian diagnosis masih menjadi tantangan besar di berbagai fasilitas kesehatan, terutama akibat ketergantungan pada *Electronic Health Records (EHR)* yang belum optimal [4]. Masalah teknis, keterbatasan kegunaan, dan

kompleksitas alur kerja EHR turut memicu kesalahan diagnosis [5]. Desain antarmuka yang kurang ramah pengguna juga memperlambat pengambilan keputusan medis, sehingga menghambat intervensi tepat waktu dan berpotensi memperburuk kondisi pasien [6].

Dalam menghadapi keterbatasan tersebut, kemajuan teknologi khususnya *Artificial Intelligence* (AI) menawarkan peluang besar untuk meningkatkan mutu layanan kesehatan melalui peningkatan akurasi diagnosis, percepatan proses medis, dan pengurangan beban kerja tenaga kesehatan [7]. Pemanfaatan *machine learning* memungkinkan analisis data medis berskala besar untuk deteksi dini dan penilaian risiko penyakit kronis, sehingga intervensi klinis dapat dilakukan lebih cepat dan efektif [8]. Selain itu, integrasi AI dengan *big data analytics* mendukung personalisasi perawatan, pengambilan keputusan real-time, dan efisiensi operasional, yang berpotensi mengoptimalkan layanan kesehatan global, selama keluaran yang dihasilkan tetap memiliki keandalan dan relevansi secara klinis [9].

Penelitian yang dilakukan oleh Anderies dan kawan-kawan [10]. Bertujuan untuk memprediksi penyakit jantung dengan memanfaatkan algoritma *machine learning* menggunakan dataset “Heart Disease Cleveland” dari Kaggle UCI Machine Learning Repository. Pada tahap pra pemrosesan, dilakukan penanganan data yang hilang, pengkodean variabel kompleks, serta pembagian data menjadi set pelatihan dan pengujian dengan rasio 80:20. Hasil evaluasi menunjukkan bahwa model Support Vector Machine (SVM) memberikan akurasi tertinggi sebesar 85% dengan nilai presisi 0,97, diikuti oleh *naive bayes* (83,33%), *logistic regression* (83,33%), *neural network* (80%) K-Nearest Neighbor (78%) dan *decision tree* (70%). Temuan ini menunjukkan bahwa SVM lebih unggul dibandingkan model lainnya dalam memprediksi penyakit jantung pada dataset tersebut.

Demikian juga penelitian yang dilakukan oleh Yang dan kawan-kawan [11] berfokus pada kemampuan klasifikasi data tabular medis, khususnya dalam tugas prediksi penyakit seperti diabetes dan kanker. Mereka memanfaatkan arsitektur LLaMA-2 yang telah melalui pra-pelatihan ulang menggunakan 12 miliar contoh tabel dari berbagai domain di Kaggle dengan metode Mask-Then-Predict serta multi-task training yang dirancang untuk mendukung klasifikasi dan regresi.

Pengujian pada dataset medis, termasuk Diabetes 130US dan dataset kesehatan lain yang mencakup diagnosis kanker, menunjukkan bahwa model ini unggul dibandingkan metode tree-based seperti XGBoost maupun model tabular berbasis deep learning. Hasil evaluasi memperlihatkan bahwa pada prediksi diabetes, model mencapai skor ROC-AUC sebesar 0,655, meningkat dari XGBoost (0,631) dan LLaMA-2 asli (0,573), sedangkan pada dataset kesehatan lainnya, akurasi juga mengalami peningkatan dengan ROC-AUC 0,868 dibandingkan XGBoost (0,854). Secara keseluruhan, peningkatan rata-rata sebesar 8,9% pada seluruh tugas klasifikasi menegaskan efektivitas pra-pelatihan LLM berbasis tabel dalam meningkatkan akurasi dan keandalan diagnosis penyakit kronis, sehingga berpotensi mendukung deteksi dini serta pengambilan keputusan klinis yang lebih tepat.

Penelitian ini bertujuan untuk mengeksplorasi pendekatan klasifikasi data medis menggunakan *classical machine learning* dan *Large Language Models* (LLM). Pendekatan ini akan diterapkan pada kumpulan data penyakit kronis, termasuk diabetes, gagal jantung, dan kanker. Data medis dalam bentuk tabular tidak hanya akan dianalisis dalam format numerik, tetapi juga diubah menjadi narasi menggunakan teknik serialisasi data. Proses serialisasi ini akan menghasilkan deskripsi tekstual yang menggambarkan kondisi pasien secara lebih komprehensif, termasuk atribut klinis seperti usia, jenis kelamin, riwayat medis, dan parameter vital.

## 1.2 Rumusan Masalah

1. Bagaimana kinerja model *machine learning* (Logistic Regression, Naive Bayes, SVM, dan XGBoost) dalam mengolah data tabular pada tiga dataset yaitu diabetes, gagal jantung, dan kanker dievaluasi menggunakan akurasi, presisi, recall, dan skor F1.
2. Bagaimana perbandingan kinerja model klasik berbasis *Natural Language Processing* (NLP) seperti **Logistic Regression, Naive Bayes, SVM, dan XGBoost** dalam melakukan klasifikasi terhadap data tabular yang telah diubah menjadi representasi teks?

3. Bagaimana kinerja large language model (LLM) seperti **LLaMA, Qwen, dan Gemini**, yang dievaluasi melalui pendekatan fine-tuningm few-shot, dan zero-shot learning dalam mengklasifikasi data medis?
4. Bagaimana hasil perbandingan antara berbagai pendekatan yang digunakan untuk mengidentifikasi kekuatan, keterbatasan, serta tingkat adaptabilitas di berbagai konteks klinis?

### 1.3 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan penelitian yang ingin dicapai berdasarkan rumusan masalah diatas, yaitu:

1. Menganalisis kinerja model machine learning (**Logistic Regression, Naive Bayes, SVM, dan XGBoost**) dalam mengolah data tabular dari tiga dataset medis (**diabetes, gagal jantung, dan kanker**) dengan menggunakan teknik klasifikasi serta evaluasi melalui metrik akurasi, presisi, recall, dan F1-score.
2. Membandingkan kinerja model klasik berbasis *Natural Language Processing* (NLP) seperti (**Logistic Regression, Naive Bayes, SVM, dan XGBoost** dalam melakukan klasifikasi terhadap data tabular yang telah dikonversi menjadi representasi teks.
3. Mengevaluasi kinerja *Large Language Model* (LLM) seperti **LLaMA, Qwen, dan Gemini** melalui pendekatan **fine-tuning, few-shot learning, dan zero-shot learning** dalam mengklasifikasikan data medis.
4. Melakukan perbandingan menyeluruh antara berbagai pendekatan yang digunakan untuk mengidentifikasi kekuatan, keterbatasan, dan tingkat adaptabilitasnya di berbagai konteks klinis.

### 1.4 Batasan Masalah

Terdapat beberapa batasan masalah dalam penelitian ini, yaitu sebagai berikut:

1. Penelitian ini terbatas pada penggunaan data penyakit Gagal Jantung 'hertfailure prediction', Diabetes 'Diabetes Dataset', dan kanker 'Cancer Prediction Dataset' sebagai sumber data utama untuk pengembangan model machine learning.
2. Penelitian ini menggunakan tiga tahap klasifikasi, yaitu klasifikasi data tabular original, klasifikasi pada dataset yang diserialisasi menggunakan metode NLP

klasik, dan klasifikasi pada dataset yang diserialisasi menggunakan large language model.

3. Penelitian ini akan memfokuskan pada eksplorasi algoritma machine learning klasik, yaitu Logistic Regression, Naive Bayes, SVM, dan XGBoost, serta Large Language Model seperti Llama, Gemini, dan Qwen.
4. Penelitian ini bertujuan menilai efektivitas ketiga pendekatan dalam klasifikasi data medis, sekaligus mengevaluasi kekuatan, keterbatasan, dan kemampuan adaptasinya di berbagai konteks klinis.

