

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah Penelitian

Diabetes Mellitus (DM) merupakan salah satu penyakit kronis yang memberikan dampak signifikan pada kesehatan masyarakat di seluruh dunia. Penyakit ini ditandai dengan tingginya kadar gula darah yang disebabkan oleh ketidakmampuan tubuh untuk memproduksi *insulin* secara memadai atau ketidakmampuan tubuh untuk menggunakan insulin secara optimal. Insulin adalah hormon penting yang berfungsi untuk membantu sel-sel tubuh menyerap *glukosa* dari darah agar dapat digunakan sebagai energi atau disimpan untuk kebutuhan energi di masa depan. Ketika fungsi insulin terganggu, *glukosa* akan menumpuk di dalam darah, yang berujung pada berbagai gangguan metabolisme. Menurut Nguyen et al. (2020), resistensi insulin memperburuk kemampuan tubuh dalam mengatur kadar gula darah, sehingga meningkatkan risiko terjadinya komplikasi serius yang berkaitan dengan gangguan metabolisme jangka panjang.

Gejala umum DM meliputi sering buang air kecil (polyuria), rasa haus berlebihan (polydipsia), dan peningkatan nafsu makan (polyphagia). Gejala-gejala ini merupakan respons tubuh terhadap kadar gula darah yang tidak stabil, yang menyebabkan terganggunya keseimbangan cairan tubuh dan metabolisme energi (Mishra & Chawla, 2021). Jika tidak dikelola dengan baik, DM dapat menyebabkan komplikasi yang mengancam jiwa, seperti penyakit jantung, stroke, kerusakan ginjal (nefropati diabetik), kerusakan saraf (neuropati diabetik), hingga kebutaan

akibat kerusakan pembuluh darah di retina (retinopati diabetik). Data dari Al-Shayea (2011) menunjukkan bahwa penderita DM memiliki risiko tinggi mengalami penyakit jantung dan pembuluh darah karena kadar gula darah yang tinggi secara kronis menyebabkan peradangan dan kerusakan pada dinding pembuluh darah. Selain itu, nefropati diabetik dilaporkan sebagai salah satu penyebab utama gagal ginjal, karena ginjal bekerja secara berlebihan untuk menyaring kelebihan gula dalam darah (American Diabetes Association, 2022). Komplikasi lain seperti retinopati diabetik juga sering kali menjadi penyebab utama gangguan penglihatan hingga kebutaan pada penderita diabetes (American Optometric Association, 2023).

Menurut data International Diabetes Federation (IDF), pada tahun 2021, lebih dari 537 juta orang di dunia hidup dengan diabetes, dan angka ini diperkirakan akan terus meningkat menjadi 643 juta pada tahun 2030. Tren ini tidak hanya menambah beban pada sistem kesehatan global, tetapi juga memengaruhi aspek sosial dan ekonomi masyarakat, terutama di negara berkembang. Salah satu tantangan utama dalam menangani diabetes adalah kurangnya akses terhadap tenaga medis ahli dan teknologi yang memadai untuk diagnosis dini. Diagnosis yang terlambat sering kali memperburuk kondisi pasien, meningkatkan risiko komplikasi, serta menambah beban biaya perawatan (Pouriyeh et al., 2017). Menurut Ghwanmeh et al. (2013), keterbatasan sumber daya medis sering kali menjadi hambatan dalam memberikan penanganan optimal bagi pasien diabetes. Oleh karena itu, diperlukan pendekatan inovatif untuk mendukung diagnosis dan manajemen diabetes agar dampaknya dapat diminimalkan.

Dalam konteks ini, Rumah Sakit Umum Bangli menjadi salah satu pusat pelayanan kesehatan yang menangani berbagai kasus diabetes. Berdasarkan data yang diambil dari rumah sakit ini, terdapat 531 data pasien yang dikumpulkan dari tahun 2021 hingga 2024. Dataset ini mencakup informasi medis penting seperti usia, jenis kelamin, tingkat gula darah, tekanan darah, serta riwayat penyakit lainnya. Analisis dataset ini dapat menjadi dasar yang kuat untuk memahami pola penyakit diabetes di daerah setempat sekaligus mengembangkan model prediktif yang dapat membantu diagnosa dini dan intervensi yang lebih tepat.

Selain itu, berdasarkan hasil observasi dan wawancara dengan pihak RSUD Bangli, diabetes menjadi salah satu penyakit yang paling sering ditangani setiap tahunnya. Banyak kasus datang dalam kondisi sudah menengah hingga berat karena tidak terdeteksi sejak dini. Hal ini memperkuat pentingnya pengembangan model klasifikasi berbasis *Machine Learning* untuk deteksi dini diabetes. Dengan demikian, pemilihan kasus diabetes dalam penelitian ini didasarkan pada tingginya prevalensi, urgensi penanganan dini, serta ketersediaan data riil dari lingkungan lokal.

Salah satu pendekatan yang menjanjikan dalam mendukung diagnosis diabetes adalah penerapan teknologi kecerdasan buatan, khususnya *Machine Learning*. Teknologi ini mampu menganalisis data dalam jumlah besar untuk mengenali pola-pola tertentu yang relevan dengan risiko diabetes. Algoritma seperti Support Vector Machine (SVM), *Decision Tree* (DT), k-nearest neighbor (KNN), dan Naive Bayes (NB) telah banyak digunakan untuk mendukung proses diagnosis di berbagai bidang medis (Ghosh et al., 2021). Namun, algoritma modern seperti *Extreme Gradient Boosting* (XGBoost) dan *Multilayer Perceptron* (MLP) menawarkan

keunggulan yang lebih signifikan, terutama dalam hal kemampuan mereka untuk menangani data besar dan kompleks.

Penelitian yang dilakukan oleh Chen et al. (2020) membandingkan performa Support Vector Machine (SVM), Neural Network, dan XGBoost dalam klasifikasi penyakit diabetes menggunakan dataset Pima Indians Diabetes Database. Hasil penelitian menunjukkan bahwa algoritma XGBoost menghasilkan akurasi tertinggi sebesar 89,2%, lebih unggul dibandingkan dengan SVM yang mencapai akurasi 86,7%, dan Neural Network dengan akurasi 88,1%. Selain itu, dari segi efisiensi waktu pelatihan, XGBoost menunjukkan keunggulan dengan durasi pelatihan sebesar 0,0021 detik, lebih cepat dibandingkan dengan Neural Network yang membutuhkan 0,0057 detik, dan SVM dengan waktu 0,0035 detik.

Penelitian lain yang dilakukan oleh Wang et al. (2022) memperluas perbandingan tersebut dengan menambahkan algoritma *Multilayer Perceptron* (MLP) ke dalam analisis. Penelitian ini menemukan bahwa MLP menghasilkan akurasi tertinggi sebesar 90,4%, mengungguli algoritma lainnya. Namun, waktu pelatihan MLP sebesar 0,0039 detik sedikit lebih lama dibandingkan dengan XGBoost.

Berdasarkan temuan dari penelitian sebelumnya, penelitian ini bertujuan untuk membandingkan kinerja algoritma XGBoost dan MLP dalam mengklasifikasikan penyakit diabetes pada dataset Rumah Sakit Umum Daerah Bangli. Fokus utama perbandingan ini adalah untuk mengevaluasi tingkat akurasi, efisiensi waktu pelatihan, serta keunggulan masing-masing algoritma dalam menangani klasifikasi penyakit diabetes. Dengan demikian, penelitian ini diharapkan dapat memberikan

wawasan mengenai algoritma yang lebih optimal dalam pemodelan klasifikasi penyakit diabetes.

1.2 Identifikasi Masalah Penelitian

1. Masih terbatasnya penelitian yang membandingkan akurasi dan efisiensi XGBoost dan MLP dalam klasifikasi penyakit diabetes pada dataset lokal, seperti rekam medis pasien RSUD Bangli.
2. Ketidakpastian dalam pemilihan algoritma terbaik untuk prediksi diabetes menimbulkan tantangan bagi tenaga medis dalam menentukan metode yang paling efektif.
3. Diperlukan studi yang membangun model klasifikasi penyakit diabetes menggunakan XGBoost dan MLP serta menganalisis perbandingan akurasi kedua algoritma untuk mendukung diagnosis yang lebih efisien.

1.3 Batasan Masalah

Agar penelitian ini lebih fokus, beberapa batasan yang diterapkan adalah sebagai berikut:

1. Dataset yang digunakan dalam penelitian diambil dari Rumah Sakit Umum Bangli dengan jumlah 531 sampel pasien yang dikumpulkan dari tahun 2021 hingga 2024.
2. Dataset mencakup 9 fitur awal: usia, jenis kelamin, kadar *glukosa*, tekanan darah, BMI, riwayat keluarga, kolesterol, aktivitas fisik, dan riwayat merokok. Namun, hanya fitur yang relevan akan digunakan dalam model

Machine Learning, yaitu Kadar *Glukosa* Darah, Tekanan Darah, Indeks Massa Tubuh (BMI), Usia, dan HbA1c

3. Tipe diabetes yang menjadi fokus penelitian adalah Diabetes Tipe 2, sesuai dengan dataset yang tersedia dari rumah sakit yang digunakan dalam penelitian ini.
4. Keputusan yang dihasilkan adalah berupa prediksi hasil klasifikasi, yang akan dibagi menjadi dua kategori: "Berisiko" dan "Tidak Berisiko."
5. Model *Machine Learning* yang digunakan adalah *Extreme Gradient Boosting* (XGBoost) dan *Multilayer Perceptron* (MLP).
6. Evaluasi model dilakukan berdasarkan akurasi, presisi, *recall*, F1-score, dan waktu pelatihan, dengan metode validasi *10-Fold Cross Validation*.

1.4 Rumusan Masalah Penelitian

Rumusan masalah penelitian ini adalah:

1. Bagaimanakah membangun model klasifikasi penyakit diabetes menggunakan algoritma *Algoritma Extreme Gradient Boosting*(XGBoost) dan *Multilayer Perceptron* (MLP)?
2. Bagaimana mengevaluasi masing-masing model klasifikasi yang dibangun?
3. Bagaimana perbandingan tingkat akurasi model klasifikasi penyakit diabetes yang menggunakan *Algoritma Extreme Gradient Boosting* (XGBoost) dan *Multilayer Perceptron* (MLP)?

1.5 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Untuk menghasilkan model klasifikasi penyakit diabetes menggunakan *Algoritma Extreme Gradient Boosting* (XGBoost) dan *Multilayer Perceptron* (MLP) dengan memanfaatkan data rekam medis pasien dari Rumah Sakit Umum Daerah (RSUD) Bangli.
2. Untuk mengevaluasi kedua model menggunakan metrik yang mencakup akurasi, presisi, *recall*, *F1-score* guna mengetahui efektivitas masing-masing algoritma dalam mendeteksi diabetes.
3. Untuk mengetahui kelebihan dan kekurangan dari masing-masing algoritma serta menentukan algoritma yang lebih optimal dalam klasifikasi penyakit diabetes berdasarkan hasil evaluasi terhadap data rekam medis pasien dari Rumah Sakit Umum Daerah (RSUD) Bangli.

1.6 Manfaat Hasil Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Manfaat Teoritis, dapat menambah pengetahuan di bidang *Machine Learning* khususnya dalam penerapan *Algoritma Extreme Gradient Boosting*(XGBoost) untuk klasifikasi penyakit biner.
2. Manfaat Praktis, menghasilkan sebuah sistem otomatis yang dapat membantu tenaga medis dalam mendeteksi penyakit diabetes, dengan tujuan mempercepat proses prediksi serta meningkatkan efisiensi dalam penanganan pasien.