

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Diabetes merupakan penyakit yang serius dan memiliki dampak signifikan terhadap kesehatan individu, dengan jumlah penderita yang terus meningkat secara pesat (Bae et al., 2022). *International Diabetes Federation* (IDF) melaporkan bahwa sebanyak 589 juta orang berusia 20 hingga 79 tahun menderita diabetes di dunia pada tahun 2024. Jumlah ini diproyeksikan meningkat menjadi 835 juta pada tahun 2050. Pada tahun yang sama, di Indonesia terdapat sekitar 20,4 juta penderita diabetes, dan angka tersebut diprediksi akan bertambah menjadi 28,6 juta pada tahun 2050 (International Diabetes Federation, 2025).

Penyakit diabetes disebabkan oleh tingginya kadar glukosa dalam darah akibat kurangnya insulin dalam tubuh atau karena tubuh tidak dapat menggunakan insulin yang diproduksi secara efektif (International Diabetes Federation, 2025). Adapun juga faktor yang berkontribusi menyebabkan diabetes, diantaranya pola hidup yang tidak sehat, obesitas, dan faktor genetic (Galicia-Garcia et al., 2020). Pola konsumsi makanan yang tidak sehat merupakan salah satu faktor terbesar yang berkontribusi terhadap penyakit diabetes (Asyikin et al., 2024). Penanganan diabetes perlu dilakukan dengan baik untuk mencegah komplikasi serius seperti gagal ginjal, gangguan saraf, kebutaan, dan masalah jantung. Penting untuk mengetahui risiko diabetes sejak dini sehingga dapat ditangani lebih awal sebelum berkembang menjadi kondisi yang sulit dikendalikan (National Institute of Diabetes and Digestive and Kidney Disease, 2023).

Seiring dengan perkembangan teknik *machine learning* dan teknologi pendukungnya, penyakit seperti diabetes mungkin dapat dideteksi risikonya, memungkinkan identifikasi dini pada manusia. Berbagai penelitian telah mengeksplorasi metode *machine learning* yang berbeda dalam mendeteksi risiko penyakit diabetes. Pradhani et al. (2025) menggunakan algoritma *Naïve Bayes* pada dataset diabetes dari Sylhet Diabetes Hospital dengan 520 sampel, menghasilkan akurasi sebesar 91% setelah proses pembersihan data. Sementara itu, Amri (2025) menggabungkan *K-Nearest Neighbors* (KNN) dengan teknik *SMOTE-ENN* pada

dataset diabetes Pima Indian, dan berhasil meningkatkan akurasi prediksi dari 69% menjadi 96%.

Zaqi (2025) membandingkan algoritma *Decision Tree*, *Random Forest*, dan *XGBoost*, serta menerapkan normalisasi *Min-Max*. *XGBoost* terbukti paling unggul dengan akurasi 99,2%. Dalam penelitian lain, Putri et al. (2021) menggunakan algoritma C4.5 pada data pasien diabetes RSUD Djasamasen Saragih dan memperoleh akurasi sebesar 90%. Sedangkan Sidiq (2025) menggunakan *AdaBoost* yang dikombinasikan dengan *SMOTE Oversampling* pada dataset diabetes dari *Kaggle*, menghasilkan akurasi sebesar 82,83%.

Suwaroyo (2023) menerapkan regresi linier menggunakan *RapidMiner* pada 200 rekam medis diabetes dan mencapai nilai *RMSE*  $0.000 \pm 0.000$ . Selain itu, Gunawan (2024) menggunakan regresi logistik pada dataset diabetes dari Puskesmas Jatibarang sebanyak 890 sampel, dengan *K-Fold Cross Validation*. Hasil evaluasi menunjukkan akurasi 82,51% dan *RMSE* sebesar 0,418. Sementara itu, Kurniati (2020) membandingkan *Naïve Bayes* dan *Support Vector Machine* (SVM) menggunakan dataset diabetes dari RSI Siti Khadijah dan mendapati bahwa SVM dengan kernel *polynomial* lebih unggul dengan akurasi 96,27%.

Hovi (2022) menggunakan SVM untuk memprediksi diabetes dan menyimpulkan bahwa model yang dilatih dengan data pasien memiliki akurasi sebesar 91%. Berdasarkan hasil penelitian yang dilakukan oleh Rohana et al. (2024) dalam memprediksi tingkat kelulusan mahasiswa menunjukkan tingkat akurasi sebesar 97,1% untuk model *Multi-Layer Perceptron* (MLP) dan 98,9% untuk model SVM. Terakhir, penelitian oleh Wijayaningrum et al. (2021) menggunakan MLP untuk memprediksi diabetes menyimpulkan bahwa arsitektur MLP dengan dua lapisan tersembunyi memiliki akurasi rata-rata sebesar 93,17%, lebih tinggi dibandingkan MLP dengan satu lapisan tersembunyi yang mencapai 92,66%.

Dari beberapa penelitian terkait, metode *machine learning* mampu mendeteksi risiko penyakit diabetes dengan performa yang baik. Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) dan *Multi-Layer Perceptron* (MLP) untuk mendeteksi diabetes. Keunggulan utama dari metode yang diusulkan penelitian ini adalah penerapan teknik *SMOTE-ENN* dalam tahap data

pre-processing untuk mengatasi *Class Imbalance* pada dataset yang digunakan, sehingga dapat meningkatkan kualitas model yang dihasilkan oleh algoritma.

## 1.2 Rumusan Masalah

Berdasarkan dari latar belakang penelitian ini, dapat disimpulkan rumusan masalah dalam penelitian ini adalah:

1. Bagaimana penerapan metode *Support Vector Machine* (SVM) dan *Multi-Layer Perceptron* (MLP) untuk memprediksi mendeteksi risiko penyakit diabetes?
2. Bagaimana perbandingan performa algoritma *Support Vector Machine* (SVM) dan *Multi-Layer Perceptron* (MLP) dalam mendeteksi risiko diabetes?

## 1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengimplementasikan metode *Support Vector Machine* (SVM) dan *Multi-Layer Perceptron* (MLP) dalam mendeteksi risiko penyakit diabetes.
2. Membandingkan performa kedua algoritma berdasarkan metrik evaluasi seperti *confusion matrix*, *accuracy*, *loss*, *precision*, *recall*, dan *F1-score*.

## 1.4 Manfaat

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan alternatif pendekatan berbasis *machine learning* dengan algoritma SVM dan MLP untuk membantu diagnosis dini diabetes secara lebih cepat dan akurat.
2. Menyediakan referensi perbandingan mengenai efektivitas algoritma SVM dan MLP dalam mendeteksi risiko penyakit diabetes