

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

*American Diabetes Association (ADA)* melaporkan bahwa ada satu orang yang terkena diabetes setiap 21 detik. Prediksi sepuluh tahun yang lalu menyatakan bahwa jumlah penderita diabetes akan mencapai 350 juta orang pada tahun 2025, namun prediksi tersebut telah jauh terlampaui. Lebih dari separuh populasi penderita diabetes di dunia berada di Asia, terutama di India, Cina, Pakistan, dan Indonesia (Yosmar et al., 2018). Diabetes adalah penyakit tidak menular yang paling umum ditemukan dalam masalah kesehatan dan penyebab kematian dan kesakitan tertinggi secara *global* (Nugroho et al., 2019). Menurut organisasi *International Diabetes Federation (IDF)*, diperkirakan setidaknya 483 juta orang pada usia 20-79 tahun di dunia mengidap diabetes pada tahun 2019, atau setara dengan tingkat prevalensi 9,3% dari total populasi pada usia yang sama. Berdasarkan jenis kelamin, *IDF* memperkirakan prevalensi diabetes pada tahun 2019 adalah 9% pada wanita dan 9,65% pada pria. Prevalensi diabetes diperkirakan akan meningkat seiring dengan bertambahnya usia penduduk menjadi 19,9% atau 112 juta orang berusia 65-79 tahun. Angka ini diprediksi akan terus meningkat hingga mencapai 578 juta di tahun 2030 dan 700 juta di tahun 2045 (Sohibul Wafa et al., 2022). Tingginya kasus diabetes pada kelompok usia produktif menyebabkan terhambatnya produktivitas (Delfina et al., 2021). Diabetes juga termasuk *silent-killer-disease*, penyebutan ini dikarenakan masih rendahnya kesadaran penderita akan penyakit tersebut sampai terjadinya komplikasi (Zulkarnaini et al., 2022).

Penyebab utama dari penyakit ini masih belum diketahui, namun para dokter berasumsi bahwa pada kondisi tersebut diduga berkaitan dengan gaya hidup dan pola makan, keduanya memainkan peran utama dalam penyakit tersebut. Penderita diabetes menghadapi resiko terkena beberapa masalah kesehatan sekunder seperti penyakit jantung, kerusakan saraf dan lain-lain (Larabi-Marie-Sainte et al., 2019). Untuk mencegah penyakit diabetes sampai pada tahap komplikasi, maka diperlukan

diagnosis sedini mungkin. Oleh karena itu penulis membuat sebuah sistem yang dapat memprediksi penyakit tersebut hanya dengan menginputkan beberapa data yang akan digunakan sebagai variabel perhitungan seperti kehamilan, gula darah, tekanan darah, lemak tubuh, insulin, berat badan, faktor keturunan dan umur. Dari perkembangan teknologi tersebut dapat membantu meringankan pekerjaan dokter dalam memprediksi penyakit diabetes (Sohibul Wafa et al., 2022). Memanfaatkan kemajuan teknologi di bidang medis yang semakin pesat dengan menggunakan kumpulan data rekam medis itu sendiri sebagai model identifikasi menggunakan teknik *data mining* (Imran et al., 2022). Salah satu teknik *data mining* yang dapat digunakan dalam memprediksi adalah klasifikasi. Tugas klasifikasi adalah menentukan hasil prediksi dengan keluaran variabel/*class* yang bernilai kategorial atau polinomial (Kotu & Deshpande, 2015 dalam Saifudin, 2018). Dengan teknik tersebut menghasilkan sebuah sistem yang dapat memprediksi penyakit diabetes. Metode klasifikasi *Support Vector Machine (SVM)* digunakan penulis pada penelitian ini dan untuk mengoptimalkan proses perhitungan metode tersebut dibutuhkan trik kernel, trik kernel berfungsi membawa data ke ruang dimensi yang lebih tinggi sehingga mempermudah proses klasifikasi (Sohibul Wafa et al., 2022). Perlu diketahui bahwa metode *SVM* adalah jenis algoritma lain dari teknik *machine learning*. Penggunaan metode *SVM* dikarenakan sifatnya yang fleksibel sehingga dapat digunakan pada banyak studi kasus.

Sebelumnya telah ada beberapa penelitian terdahulu yang berkaitan dengan penelitian ini, penelitian (Anam et al., 2023) tentang analisis sentimen pembelajaran online pada media sosial *twitter* dengan menggunakan metode *SVM* pada kernel *linear*. Dengan menggunakan perbandingan 80%:20% antara data latih dan data uji didapatkan hasil akurasi sebesar 90% pada pelabelan otomatis menggunakan portal *Drone Emprit Academy*, sedangkan pada pelabelan menggunakan metode *lexicon-based*, hasil akurasi meningkat menjadi 95%. Berikutnya penelitian dari (Afifah et al., 2023) yang melakukan perbandingan antara kernel *Linear* dan *Radial Basis Function (RBF)* pada metode *Support Vector Machine (SVM)* dalam memprediksi kasus serangan *Ransomware Locker*. Hasil penelitian tersebut menunjukkan bahwa kernel *RBF* lebih unggul dengan akurasi 93.27% pada *AUC Train* dan

93.41% pada *AUC Test*. Penelitian lainnya dari (Fitriyah et al., 2020), mengenai analisis sentimen pada gojek di media sosial *twitter* menggunakan metode *Support Vector Machine (SVM)* dengan kernel *Linear* dan *Radial Basis Function (RBF)*. Akurasi tertinggi dihasilkan oleh metode *SVM* pada kernel *RBF* dengan pelabelan manual atau *sentiment scoring* sebesar 79.19%.

Mengingat betapa bermanfaatnya inovasi dari sebuah teknologi yang dapat memprediksi penyakit diabetes, maka penulis mengeksekusi solusi tersebut dengan menerapkan metode klasifikasi *Support Vector Machine (SVM)* dengan kernel *polynomial*. Penggunaan metode *SVM* dikarenakan metode tersebut dapat mengklasifikasikan *margin* setiap kelas secara optimal dan sangat cepat dalam proses perhitungan komputasi yang kompleks (Purnawansyah et al., 2023). Teknik ini juga dapat menyelesaikan masalah klasifikasi dan regresi secara *linear* ataupun *nonlinear* (Sudirman et al., 2024). Penelitian ini bertujuan untuk membantu para masyarakat dalam melakukan prediksi penyakit diabetes secara cepat dan akurat dan tentunya lebih hemat biaya walaupun dibutuhkan nilai inputan yang bersifat medis sehingga membutuhkan peran dokter juga disana. Namun sistem ini dapat menjadi alternatif pemotongan pengeluaran anggaran setiap individu masyarakat.

## 1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, didapat rumusan masalah sebagai berikut :

1. Seberapa efektif penggunaan metode *Support Vector Machine (SVM)* pada kernel *polynomial* dalam melakukan prediksi penyakit diabetes?
2. Seberapa berpengaruh variabel parameter perhitungan metode *Support Vector Machine (SVM)* dan penggunaan rasio perbandingan data dalam menentukan hasil akurasi?

## 1.3 Batasan Masalah

Agar penelitian tidak keluar dari pokok permasalahan yang dirumuskan, maka ruang lingkup permasalahan dibatasi pada :

1. Sistem berfokus pada implementasi metode *Support Vector Machine (SVM)* dengan *Polynomial Kernel* serta menghitung nilai akurasi.
2. Rentang waktu dataset yang digunakan pada tahun 2022-2024.
3. Sistem dapat melakukan prediksi penyakit diabetes dengan menggunakan nilai variabel yang telah diinputkan pada sistem oleh pengguna.
4. Sistem prediksi penyakit diabetes dapat diakses oleh umum, dengan syarat telah memiliki akun terdaftar kepada sistem untuk *login* terlebih dahulu.

#### **1.4 Tujuan**

Adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Mempermudah masyarakat dalam melakukan identifikasi diri akan penyakit diabetes yang mungkin menimpa mereka.
2. Menghitung tingkat akurasi metode klasifikasi *Support Vector Machine (SVM)* pada kernel *polynomial*.

#### **1.5 Manfaat**

Adapun manfaat dari penelitian ini sebagai berikut :

1. Memahami penerapan metode *Support Vector Machine (SVM)* pada *kernel polynomial* serta mengetahui tingkat akurasi metode tersebut.
2. Mendapatkan hasil diagnosis akan terkena atau tidaknya terhadap penyakit diabetes lebih awal guna terhindar dari komplikasi.