

BAB I

PENDAHULUAN

1.1.Latar Belakang

Diabetes atau istilah lainnya yaitu (*Diabetes Melitus*) merupakan sebuah penyakit yang dimana kandungan gula darah dalam tubuh tidak dapat diserap atau diolah oleh tubuh. Maka kondisi normal yang ada pada seseorang secara otomatis dapat menghasilkan insulin yang berfungsi untuk memastikan kadar gula tetap berada diatas ambang batas yang normal. Ketika tubuh tidak cukup menghasilkan *hormone insulin* atau sel tidak dapat merespon dari insulin, maka gula darah akan menjadi tinggi sehingga penyakit dari diabetes itupun akan terjadi. Gejala-gejala klasik yang terjadi secara umum adalah hasrat ingin buang air kecil terlalu sering, rasa haus yang tidak henti-henti, dan tingginya rasa lapar [1].

Data terbaru dari *International Diabetes Federation (IDF)* mengemukakan bahwa Indonesia terdapat pada urutan ke-7 dari 10 negara dengan jumlah pasien diabetes tertinggi [2]. Persentase pasien pengidap penyakit diabetes di Indonesia mencapai 6,2%, yaitu sebanyak 10,8 juta orang menderita penyakit diabetes per tahun 2020 di Indonesia. Untuk penderita diabetes tersebut cenderung lebih tinggi berjenis kelamin laki-laki dibandingkan dengan Wanita [2]. Hal ini disebabkan perubahan pola dan gaya hidup warga sekitar yang ingin menjadi lebih sehat.

Pada banyak kasus penderita diabetes di Indonesia dan seluruh dunia, diperlukan tindakan sebagai solusi dari permasalahan dari penyakit diabetes itu sendiri. Hal tersebut pasti perlu dilakukan dengan keputusan medis yang dibuat dengan pengalaman dan penalaran secara medis. Prediksi itu sendiri dapat

dilakukan dengan berbagai cara, dapat dengan melakukan cara memanfaatkan beberapa data pasien penderita penyakit diabetes yang telah tersimpan dalam *dataset* agar dibuat pola dengan menggunakan klasifikasi dari bahasa mesin sehingga hasil yang akan didapat tidak jauh beda dan tepat dapat dihindari. Sebuah inovasi dan kontribusi dalam perencanaan dimasa yang akan datang menggunakan teknologi algoritma yang tepat dari prediksi yang akurat menggunakan algoritma *K-Nearest Neighbor (KNN)*.

Pada penelitian ini, berfokus pada algoritma K-Nearest Neighbor (KNN), yaitu metode yang digunakan untuk mengklasifikasikan sekumpulan data berdasarkan pembelajaran data yang terdapat pada data training yang ada. Melalui pengenalan pola algoritma tersebut mampu melakukan pemodelan yang akan dibuat, di samping kelebihan terdapat kekurangannya dalam algoritmanya itu sendiri yaitu salah satunya tidak dapat menangani nilai yang hilang (*missing value*) secara implisit. Selain menggunakan algoritma KNN, pada penelitian ini juga menggunakan suatu metode, yaitu metode machine learning. Machine learning merupakan mesin yang dapat bekerja sama dengan cara melakukan pendekatan kepada manusia, yang berguna menyelesaikan akar permasalahan dengan cara yang terukur [3].

Dari penelitian ini akan diketahui bahwa algoritma KNN dapat digunakan untuk memprediksi penyakit Diabetes Mellitus atau tidak dengan menggunakan algoritma KNN yang diharapkan dapat memberikan ketepatan dalam memprediksi penyakit Diabetes Mellitus berdasarkan data yang telah tersedia dari dataset yang akan digunakan dalam pengolahan dan pembuatan model.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka perlu rumusan masalah dalam penelitian ini yaitu :

1. Apakah algoritma K-Nearest Neighbor (K-NN) dapat memprediksi penyakit *Diabetes Mellitus*?
2. Berapa tingkat akurasi yang didapat dari algoritma K-Nearest Neighbor (K-NN) dalam memprediksi penyakit *Diabetes Mellitus*?

1.3. Batasan Masalah

Batasan masalah dalam Skripsi ini adalah:

1. Pendeteksi dari penyakit ini berbasis web, dengan menggunakan Bahasa pemrograman Python dan HTML untuk pembuatan sistemnya.
2. Metode dan algoritma yang digunakan untuk memprediksi penyakit *Diabetes* menggunakan algoritma K-Nearest Neighbor (K-NN).
3. Metode untuk mengevaluasi akurasi hasil klasifikasi menggunakan Confusion Matrix.
4. Hasil dari prediksi awal penyakit *Diabetes Mellitus* adalah dengan menampilkan nilai positif dan negatif.
5. Jumlah dataset digunakan 768 data penyakit diabetes.

1.4. Urgensi Penelitian

Penelitian mengenai pembuatan program aplikasi untuk mendeteksi penyakit *Diabetes Mellitus* dengan pengimplementasian dan telah dikembangkan menjadi sistem web. Dibutuhkan sistem pendukung keputusan yang dapat membantu pengguna yaitu masyarakat umum ataupun petugas medis pada puskesmas ataupun rumah sakit agar dapat memprediksi secara cepat pasien tersebut didiagnosa menderita penyakit diabetes atau tidak. Pada penelitian ini menggunakan suatu teknik dalam *Machine Learning* yaitu *K-Nearest Neighbor* (K-NN) dengan menggunakan 8 variabel yaitu, jumlah kehamilan yang pernah diderita pasien, kadar glukosa, tekanan darah, lemak tubuh, kadar insulin, berat badan (BMI), dan riwayat penyakit diabetes dari keluarga pasien. Dikarenakan

pada K-NN ini merupakan salah satu metode yang digunakan untuk mengklasifikasikan objek berdasarkan data latih dengan jarak yang terdekat. Pada sistem ini diharapkan dapat menghasilkan tingkat keakuratan yang jauh lebih tinggi, sehingga diharapkan dapat mencegah kemungkinan terburuk yang akan terjadi akibat penyakit *Diabetes Mellitus*.

1.5.Maksud dan Tujuan Penelitian

Tujuan dari pembuatan Skripsi ini adalah:

1. Dapat menerapkan algoritma K-Nearest Neighbor agar dapat digunakan untuk memprediksi penyakit Diabetes.
2. Dapat memanfaatkan dataset untuk memprediksi penyakit Diabetes.
3. Mencari nilai akurasi algoritma K-Nearest Neighbor dalam memprediksi awal penyakit Diabetes.

1.6.Manfaat Penulisan

a. Bagi Masyarakat

Dapat memudahkan masyarakat untuk memprediksi penyakit Diabetes, dan kemudian dapat melakukan pemeriksaan sedini mungkin.

b. Bagi Ilmu Pengetahuan

1. Dapat mengimplementasi algoritma K-Nearest Neighbor dalam pemrograman *Python*.
2. Mengetahui cara kerja algoritma K-Nearest Neighbor dalam memprediksi suatu kondisi.

1.7. Metode Penelitian

Metode penelitian yang akan digunakan dalam pembuatan skripsi ini meliputi:

1.7.1 Metode Pengumpulan Data

Metode pengumpulan data yang akan dilakukan untuk penelitian ini, yaitu:

1. Mencari referensi atau bahan baca tentang algoritma *K-Nearest Neighbor*, *Accuracy Score*, *Confusion Matrix* dan *Diabetes Mellitus*.
2. Mendapatkan dataset publik tentang gejala awal penderita penyakit diabetes. Berasal dari *Kaggle*, *Pima Indians Diabetes Database*.

1.7.2. Metode Analisis

1. Metode *K-Nearest Neighbor*
Membahas langkah-langkah yang akan dilakukan dengan menggunakan metode algoritma *K-Nearest Neighbor*.
2. Menghitung akurasi dari pengujian data
Untuk menguji akurasi dari prediksi data menggunakan metode *Confusion Matrix*.

1.7.3. Metode Perancangan

Merancang alur dari sistem yang akan dibuat menggunakan flowchart, membuat Data Flow Diagram yaitu untuk mengetahui alur penggunaan dari sistem yang akan dibuat.

1.7.4. Metode Implementasi

Implementasi algoritma *K-Nearest Neighbor* menggunakan bahasa Python. Di implementasikan dengan 2 cara yaitu, pembuatan model dengan jupyter notebook kemudian pembuatan sistem dengan web,

menggunakan HTML dan CSS.

1.7.5. Metode Pengujian

Menginputkan data uji dan data latih yang telah digabungkan secara acak dan sistem akan menampilkan hasil proses klasifikasi yang menyebutkan apakah penderita terkena diabetes atau tidak.

1.8. Sistematika Penulisan

Sistematika dalam penulisan Skripsi ini dibagi menjadi lima bab, antara lain sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan mengenai latar belakang suatu masalah, perumusan masalah, batasan masalah, tujuan, manfaat, urgensi, serta sistematika dalam penulisan penelitian.

BAB II LANDASAN TEORI

Pada bab ini menjelaskan mengenai teori dasar yang digunakan untuk menganalisis penelitian. Meliputi tinjauan pustaka penyakit diabetes, algoritma K-Nearest Neighbor, Python, Machine Learning, dan Confusion Matrix.

BAB III ANALISIS DAN PERANCANGAN

Pada bab ini menjelaskan mengenai perancangan dan pemodelan serta cara kerja yang dilakukan dalam pembuatan aplikasi pendeteksi penyakit diabetes berbasis website menggunakan algoritma K-Nearest Neighbor dan Machine Learning.