

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Kesehatan merupakan hal yang selalu diinginkan oleh setiap manusia, karena dengan kesehatan segala kegiatan dapat dilakukan dengan maksimal. Sehingga bukan hal yang tidak mungkin jika seseorang mengorbankan segala hal untuk menjaga kesehatannya. Diabetes atau biasa disebut penyakit gula dapat dialami oleh siapapun tidak hanya dari faktor keturunan, selain itu salah satu faktor penyebab diabetes yaitu terlalu banyak mengonsumsi gula berkepanjangan. Gula itu sendiri memiliki peran yang penting yaitu sumber penyimpanan energi di dalam tubuh, tetapi jika mengonsumsi secara berlebihan menyebabkan kekebalan pada hormon insulin dan gula darah tubuh menjadi terlalu tinggi. Gula darah tubuh yang tinggi dan tidak terkontrol menyebabkan penyakit diabetes.

Perkembangan teknologi yang pesat, khususnya pada sistem pengujian atau pengidentifikasian yang sering digunakan sebagai objek riset dan pengembangan dengan teknologi komputer. Mengidentifikasi suatu objek dengan sistem ini menjadi salah satu riset yang populer. K-Nearest Neighbors merupakan salah satu metode klasifikasi pada data mining dan juga menjadi algoritma *supervised learning* pada *machine learning* dengan metode pengelompokkan suatu hal berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Algoritma K-Nearest Neighbor mengelompokkan data baru yang belum diketahui kelasnya dengan memilih data sejumlah k yang letaknya paling dekat dari data baru. Metode KNN menggunakan nilai k sebagai 2 parameter untuk menghitung jumlah klasifikasi yang jaraknya terdekat dengan objek dan mengambil mayoritas klasifikasi untuk dipilih sebagai keputusan.

Data tentang penyakit diabetes yang diperoleh dari dataset *National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK)* yang dibagikan oleh Mehmet Akturk pada situs *Kaggle* terlihat bahwa terdapat atribut umum (*Pregnancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, Diabetes Pedigree Function, Age*) dan kelas (Outcome). Pada data yang ada penulis

mencoba melakukan penelitian menggunakan metode KNN untuk menghitung akurasi berdasarkan *N-neighbor* yang digunakan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka dirumuskan suatu masalah yang akan dibahas pada penelitian ini adalah menguji akurasi dari metode KNN berdasarkan variabel pada dataset diabetes berdasarkan *N-neighbor* yang digunakan.

1.3 Batasan Masalah

Agar penelitian lebih terarah dan sesuai dengan tujuan penelitian, maka penulis membatasi permasalahan sebagai berikut :

- a. Dataset diabetes diperoleh dari situs kaggle jumlah total 768 data, yang terdiri dari 500 data negatif penyakit diabetes dan 268 data positif penyakit diabetes. Data dari dataset diabetes adalah perempuan berusia minimal 21 tahun dari keturunan Pima India.
- b. Variabel independent yang digunakan adalah *Pregnancies, Glucose, BMI, Age* dan variabel dependennya adalah *outcome*.
- c. Bahasa pemrograman yang digunakan adalah *Python*.
- d. Penelitian hanya dilakukan untuk mengukur seberapa akurat algoritma *K-Nearest Neighbor* (KNN) berdasarkan *N-neighbor* yang digunakan.

1.4 Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah diatas tujuan yang ingin diraih dalam pembuatan laporan skripsi ini adalah untuk mengetahui perbandingan akurasi metode *K-Nearest Neighbor* (KNN) pada variabel dataset diabetes yang dipilih berdasarkan *N-neighbor* yang digunakan.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah

1. Hasil penelitian ini dapat menjadi acuan penelitian lebih lanjut dengan melakukan perbandingan lainnya menggunakan metode *K-Nearest Neighbor* (KNN).

2. Dapat menjadi metode alternatif suatu pihak dalam perancangan sebuah sistem untuk mengidentifikasi penyakit diabetes menggunakan algoritma *K-Nearest Neighbor* (KNN).

1.6 Sistematika Penulisan

Penulisan skripsi ini terdapat lima bab yang tersusun secara sistematis guna mempermudah pembaca. Sistematika penulisan skripsi ini sebagai berikut:

BAB I PENDAHULUAN

Pada Bab I menjelaskan tentang latar belakang suatu masalah, rumusan masalah, Batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini berisi uraian tentang teori – teori dari penelitian sebelumnya yang berhubungan dengan penelitian yang dibuat.

BAB III METODE PENELITIAN

Berisi penjelasan tentang metode yang digunakan untuk memahami dan mengeksplorasi obyek penelitian, pengumpulan data, masalah yang ada pada obyek yang diteliti, dan rancangan alur penelitian pada skripsi.

BAB IV PEMBAHASAN

Bab ini menjelaskan mengenai rancangan, implementasi serta evaluasi dari rancangan. Selanjutnya alur pengerjaan proyek, metode testing, hingga hasil akhir penelitian dan pembahasan analisis hasil akhir penelitian, termasuk pembahasan hasil-hasil uji coba (*testing*).

BAB V PENUTUP

Bab ini berisikan tentang kesimpulan dan saran dari penelitian.