

# BAB I

## PENDAHULUAN

### 1.1 LATAR BELAKANG

Sesuatu yang sangat diharapkan dalam mengumpulkan informasi adalah menyampaikan informasi secara menyeluruh. Meskipun demikian, bila ada informasi yang kurang atau informasi yang hilang di dalamnya, alasannya harus dijelaskan terlebih dahulu. Sebagai aturan umum, informasi yang tidak memadai terjadi karena faktor-faktor tertentu yang membuat informasi terfragmentasi. Kumpulan informasi adalah kumpulan data yang disimpan cukup lama dan berubah menjadi tumpukan informasi yang sangat besar. Jarang terjadi pengumpulan informasi yang digunakan dalam ujian yang diperkenalkan belum selesai. Tentunya hal ini berdampak pada siklus yang akan dilakukan nantinya. Indeks informasi pada umumnya mempunyai informasi total, namun dalam banyak kasus terlihat adanya informasi yang hilang (missing value) dalam kumpulan informasi. Dalam penelitian, adanya missing value dalam kumpulan data akan menjadi masalah penting yang perlu diatasi.

KNN adalah metodologi yang digunakan untuk mengenali objek berdasarkan data tertentu yang jarak terdekatnya dengan artikel. KNN bekerja dengan memastikan jarak antara informasi baru dan informasi (informasi uji). Meskipun keteraturan adalah metode yang terlibat dalam pencarian kepribadian khas antara berbagai elemen dan mengkaraktisasinya ke dalam kelas yang sesuai, pengelompokan informasi dalam jumlah besar dan kompleks bila dilakukan secara fisik akan menyusahakan dan menyita sebagian besar waktu. Hal ini dapat dilakukan dengan mengembangkan perangkat ketertiban karena dengan instrumen pengelompokan informasi yang sangat besar dan kompleks, pengiriman dapat dilakukan dengan mudah dan cepat sesuai dengan kebutuhan masing-masing kelas. Salah satu strategi karakterisasi yang dapat digunakan untuk mengatasi masalah kualitas yang hilang dalam perhitungan KNN.

Perhitungan KNN merupakan suatu kerangka yang memanfaatkan perhitungan pembelajaran dan rencana untuk menemukan desain informasi baru dengan menghubungkan desain informasi yang sudah ada. Perhitungan NN ada dua macam, yaitu INN dan KNNI, INN atau Neighbor merupakan metodologi yang

memanfaatkan pengaturan informasi K terdekat dengan  $K > 1$ . Pada tahun 2019, Dese Narfa Firmasyah dkk. melakukan penelitian yang temuannya menunjukkan bahwa algoritma self-organizing map (MOS) berhasil membentuk kelompok (cluster) dengan siluet tertinggi 0,351 ketika konfigurasi optimal epoch penelitian adalah 160 dan jumlah neuron  $2 \times 2$ . Terlebih lagi, sekitar waktu yang sama tahun 2019, kembali dilakukan kajian oleh Iman Jihad Fadillah dan Siti Muchlisson yang bertajuk korelasi teknik Hot-Deck Ascription dan strategi KNNI dalam mengelola kualitas yang hilang, yakni perhitungan KNNI dapat menampilkan kembali informasi yang hilang dan sebagai sejauh presisi perhitungan KNNI menghasilkan ketepatan penilai. lebih tepat dan dapat diandalkan.

Sistem ini dapat menghitung nilai variabilitas glikemik. MAGE adalah variabilitas glikemik yang dapat dihitung menggunakan sistem ini. MAGE biasanya ditampilkan sebagai sistem pemantauan glukosa darah secara kontinyu dengan menggunakan alat yang disebut CGM (Continuous Glucose Monitoring). Pengukuran data yang diperoleh masih berupa data diskrit dan harus diubah menjadi pendekatan teknik interpolasi data kontinyu (ICGM / Interpolated Continue Glucose Monitoring), sehingga model data yang dihasilkan sama dengan CGM. Kadar glukosa pada pasien diabetes diukur menggunakan glukometer Accu Check Performa Nano. Terdiri dari sensor (strip) yang berfungsi mengubah glukosa menjadi sinyal atau besaran listrik dan ditampilkan di monitor dalam bentuk data numerik (diskrit) yang dapat disimpan dalam memori yang terintegrasi dengan alat.

Berdasarkan hasil penelitian sebelumnya yang dilakukan oleh Y&C, dimana meneliti terkait comparative analysis of Classification Methods of KNN and naïve Bayes to Determine Stress level of junior High School Students. hasil penelitiannya enunjukkan bahwa KNN dengan  $K=3$  yakni 86.61% dan untuk metode Naïve Bayes diperoleh nilai dengan 87.40% di tungkai bahkan yan mengalami penderitaan bisa saja tidak mendapatkan adanya perubahan.

Keunikan yang terjadi berdasarkan laporan The Worldwide Diabetes League, peningkatan jumlah penderita diabetes pada tahun 2003 hingga 2025 akan meningkat. Indonesia sendiri berada di peringkat kelima dengan jumlah penderita diabetes yang berjumlah 12,9 juta orang dan diperkirakan akan terus naik ke peringkat ketiga dunia pada tahun 2025. Ada berbagai strategi pengelompokan

dalam pembelajaran teregulasi di AI, antara lain KNN Gullible Bag, Backing Mesin Vektor, Organisasi Otak, Pengklasifikasi Hutan Tidak Beraturan, terdapat Pengklasifikasi Lift dan Ujian Diskriminan Kuadrat. Dalam eksplorasi ini, strategi KNN digunakan untuk siklus penataan. Strategi karakterisasi menggunakan KNN merupakan teknik yang akan segera dibuat untuk mengelompokkan diabetes. Strategi KNN memiliki beberapa manfaat, khususnya, persiapannya sangat cepat, mudah, dan efektif dengan informasi persiapan yang sangat besar. Sedangkan kekurangan dari KNN sendiri adalah nilai predisposisi K yang tidak sepenuhnya menetap, komputasi yang membingungkan, keterbatasan memori dan juga dapat diakali dengan sifat-sifat yang tidak penting. Perhitungan KNN menggunakan informasi yang ada pada data hasil pengukuran atau data real, sehingga dapat diperoleh hasil dari rencana tersebut yang akan dipikirkan dan dapat digunakan untuk kebutuhan klinis dan pengelompokan pada diabetes.

Berdasarkan fenomena diatas maka penulis tertarik melakukan penelitian dengan judul **“Analisa Kesalahan Penginputan Data Menggunakan Metode KNN Untuk Data Penyakit Diabetes”**

## **1.2 RUMUSAN MASALAH**

Dilihat dari latar belakang yang telah dijelaskan diatas, sehingga rumusan masalah yang dapat ditraik pada penelitian ini adalah sebagai berikut :

1. Bagaimana menyelesaikan masalah missing value pada data penyakit diabetes menggunakan metode KNN ?
2. Bagaimana menghitung nilai “K” paling optimal pada KNN saat implementasi missing value dilakukan ?

## **1.3 TUJUAN PENELITIAN**

Dari rumusan masalah dapat ditarik pula tujuan masalah yang akan dimabil pada penelitian ini ,:

1. Menganalisis efektivitas metode KNN dalam imputasi missing value pada data penyakit Diabetes.
2. Memperoleh nilai “ K “ paling optimal berdasarkan hasil penelitian.
3. Melengkapi nilai missing value pada data penyakit diabetes.

#### 1.4 BATASAN PENELITIAN

Agar pembahasan dalam penelitian yang dilakukan lebih terfokus, maka penulis membatasi masalah sebagai berikut:

1. Menggunakan perbaikan missing value dengan metode imputasi KNN.
2. Nilai ' K ' atau tetangga yang digunakan pada penelitian ini adalah ' K '=3, 'K=5, 'K= 7, 'K = 7, ' K=9, dan ' K=15.
3. Melengkapi nilai missing value pada atribut yang tidak memiliki nilai.

#### 1.5 MANFAAT PENELITIAN

Adapun manfaat dari penelitian ini adalah sebagai berikut;

- a. Bagi penulis

Manfaat penelitian ini bagi penulis ialah sebagai penerapan metode dan ilmu yang diperoleh dan melatih untuk menganalisis sesuatu yang sudah ada juga untuk mencari penyelesaiannya

- b. Bagi pembaca

Selain bagi penulis penelitian memiliki manfaat bagi pembaca yakni sebagai bahan wawasan dan pengetahuan juga sebagai bahan untuk membandingkan juga sebagai sumber yang akan digunakan pada kajian yang sama.