

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Diabetes mellitus ialah penyakit yang diindikasikan kenaikan kadar gula dalam darah. Penyakit ini disebabkan karena tubuh tidak mampu menggunakan serta memproduksi hormon insulin secara efektif. Insulin adalah hormon yang dibentuk oleh pankreas yang berfungsi dalam mengatur kadar gula darah pada tubuh. Diabetes mellitus ditandai dengan tingkat glukosa dalam darah yang tinggi pada periode yang lama. Kadar glukosa pada darah dalam dua jam setelah mengonsumsi dan meminum cairan yang memiliki kandungan gula dan karbohidrat lain tidak melebihi 140 mg/dL (Lesmana, 2012). Penyakit diabetes mellitus ini dapat mengakibatkan berbagai macam komplikasi kronis contohnya stroke, serangan jantung, kebutaan serta amputasi (Rudy Ariyanto dkk., 2020).

Diperkirakan oleh Organisasi Kesehatan Dunia (WHO), banyaknya pengidap penyakit diabetes melitus di Indonesia nantinya terus mengalami kenaikan dengan signifikan sampai 21,3 juta jiwa di tahun 2030. Menurut Riset Kesehatan Dasar 2018 (Riskesdas) jumlah keseluruhan kasus atau prevalensi diabetes melitus di Indonesia untuk usia 15 tahun ialah senilai 2%. Prevalensi orang penderita diabetes melitus di Indonesia mengindikasikan peningkatan yakni dari 5,7% di tahun 2007 naik menjadi 6,9% di tahun 2013. Di Provinsi Bali sekitar 52.251 orang berdasarkan data kesehatan Provinsi Bali merupakan penderita penyakit diabetes mellitus, sementara di Kabupaten Buleleng menurut data dari Dinas Kesehatan Kabupaten Buleleng tahun 2021 dimana penyakit diabetes mellitus menduduki peringkat kedua

dari 10 peringkat teratas penyakit yang ditemukan di Kabupaten Buleleng yaitu dengan jumlah kasus sekitar 15.089 orang. Dari jumlah tersebut terdapat 7.102 kasus dengan jenis kelamin laki-laki serta 7.987 kasus dengan berjenis kelamin perempuan. Kecamatan Buleleng merupakan wilayah di Kabupaten tersebut yang memiliki jumlah kasus penderita diabetes mellitus terbanyak yaitu sekitar 1.740 kasus, berdasarkan data dari seksi pencegahan dan penanggulangan penyakit tidak menular Dinas Kesehatan Kabupaten Buleleng.

Mestinya penyakit diabetes mellitus dapat dideteksi sedini mungkin. Namun kenyataannya penyakit diabetes mellitus umumnya dapat di deteksi setelah adanya gejala atau tanda-tanda klinis yang muncul. Diabetes mellitus mempunyai gejala umum contohnya lebih banyak kencing, rasa haus yang berlebihan, infeksi berulang, dan gejala lainnya. Penting untuk menyadari bahwa tidak semua orang dengan diabetes mellitus akan mengalami gejala-gejala seperti itu. Beberapa orang mungkin memiliki diabetes mellitus tanpa menunjukkan gejala sama sekali. Maka, sangat penting untuk mendeteksi secara dini dalam upaya mencegah komplikasi yang lebih serius. Untuk mendeteksi diabetes mellitus, tes darah umumnya digunakan. Beberapa tes yang umum digunakan seperti pemeriksaan gula darah puasa, tes toleransi glukosa oral (OGTT), serta ada juga tes A1C (HbA1c) yang memberikan gambaran tentang rerata kadar gula darah pada periode 2 hingga 3 bulan terakhir.

Teknologi informasi di Indonesia dewasa ini sudah mengalami perkembangan yang masif sehingga diaplikasikan di berbagai lembaga dan juga instansi yang membutuhkan pemrosesan data dalam skala besar. Salah satu bidang yang memerlukan pemrosesan data yang dalam skala besar adalah di bidang medis.

Dalam hal memprediksi penyakit di bidang medis khususnya dalam ranah ilmu pengetahuan ataupun teknologi yang sudah berkembang pesat, oleh karena itu penerapan teknologi di bidang medis hanya dilakukan pada fasilitas pelayanan kesehatan yang memenuhi persyaratan dan hanya diaplikasikan oleh sumber daya manusia (SDM) yang mempunyai kewenangan dan berkompeten, sehingga nantinya dapat memberikan perlindungan yang sesuai bagi masyarakat. Hal ini, diperlukan untuk mencegah penyebaran penyakit yang membahayakan lingkungan sekitar dan juga untuk menghindari risiko lainnya.

Dalam melakukan proses data dengan jumlah besar diperlukan teknik penggalian data seperti *data mining* yakni proses penemuan pola dan informasi yang penting dari kumpulan data yang besar. Ini melibatkan penggunaan teknik komputasional untuk mengidentifikasi pola yang tersembunyi, hubungan, dan tren yang bisa dimanfaatkan dalam mengambil keputusan dan prediksi. Tujuan dari *data mining* ialah untuk mengeksplorasi dan menganalisis data dalam rangka mengungkapkan wawasan baru dan pengetahuan yang dapat mendukung pengambilan keputusan yang lebih baik. Adanya teknik *data mining* dapat membantu untuk melakukan penelitian dari suatu penyakit salah satunya penyakit diabetes mellitus melalui variabel data seperti jenis kelamin, riwayat penyakit tidak menular dalam diri sendiri, riwayat penyakit tidak menular dalam keluarga, berat badan, tinggi badan, *body mass index* (BMI), usia, lingkar perut, pemeriksaan gula, dan diagnosis. Hal ini akan memudahkan dalam mengidentifikasi pola keputusan terkait penyakit diabetes mellitus karena faktor-faktor yang mempengaruhi hasil suatu diagnosis sudah diketahui. Salah satu algoritma yang dipergunakan untuk mengidentifikasi keputusan tersebut ialah mempergunakan algoritma klasifikasi.

Dalam penelitian sebelumnya yang dilakukan oleh (Robbani dkk., 2022) dengan judul “Klasifikasi Penyakit Diabetes Mellitus Berdasarkan Faktor-Faktor Penyebab Diabetes Menggunakan Algoritma C4.5”, digunakan sumber data dari *National Institute of Diabetes and Digestive and Kidney Diseases*. Data tersebut berasal dari *Prima Indians Diabetes* yang diambil dari sumber terbuka *Kaggle*, terdiri dari 768 data dengan 9 variabel. Penelitian ini menggunakan 9 variabel, yakni usia, riwayat keluarga, BMI, insulin, ketebalan kulit, tekanan darah, kadar glukosa, banyaknya kehamilan serta hasil akhir (*outcome*). Pengukuran akurasi dilakukan dengan membandingkan data latih serta data uji dalam perbandingan 80:20, dan hasilnya menunjukkan tingkat keakuratan sebesar 76% dalam klasifikasi.

Dalam penentuan status penyakit diabetes mellitus dibutuhkan algoritma yang tepat untuk melakukan proses pengidentifikasian. Peneliti menggunakan algoritma klasifikasi *data mining* yakni *K-Nearest Neighbor* dalam melakukan identifikasi penyakit diabetes mellitus di Kecamatan Buleleng. Hal ini bertujuan untuk memberikan pandangan yang lebih jelas dalam memprediksi penyakit diabetes mellitus di Kecamatan Buleleng.

Algoritma *K-Nearest Neighbor* ialah suatu algoritma klasifikasi pada *data mining* yang bekerja sesuai dengan “*nearest neighbor*”. Pada dasarnya, *K-Nearest Neighbor* bekerja dengan mengukur jarak diantara data yang nantinya diprediksi dengan data pada setiap kelas yang ada, kemudian data dengan jarak paling dekat dengan data yang nantinya diprediksi dijadikan sebagai acuan untuk memprediksi kelas data. Algoritma *K-Nearest Neighbor* dipergunakan dalam kasus klasifikasi data yang terdiri dari beberapa variabel atau atribut. Kelebihannya ialah mempunyai

sedikit *hyperparameter* karena algoritma *K-Nearest Neighbor* hanya memerlukan nilai  $K$  serta metrik jarak lebih kecil dibanding algoritma *machine learning* yang lain, serta memiliki tingkat akurasi yang dapat diterima.

Berlandaskan pada masalah yang sudah dijelaskan tersebut, peneliti melaksanakan “**Identifikasi Penyakit Diabetes Mellitus di Kecamatan Buleleng dengan Menerapkan Algoritma *K-Nearest Neighbor*”**”.

## 1.2 Rumusan Masalah

Sesuai latar belakang yang telah dipaparkan, bisa dibentuk rumusan permasalahan berikut.

1. Bagaimana hasil penerapan algoritma *K-Nearest Neighbor* untuk mengklasifikasi penyakit diabetes mellitus di Kecamatan Buleleng?
2. Bagaimana desain dari aplikasi yang diterapkan untuk mengklasifikasi penyakit diabetes mellitus di Kecamatan Buleleng?
3. Bagaimana implementasi aplikasi yang menerapkan algoritma *K-Nearest Neighbor* untuk mengklasifikasi penyakit diabetes mellitus di Kecamatan Buleleng?
4. Bagaimana keakurasian aplikasi yang menerapkan algoritma *K-Nearest Neighbor* dalam klasifikasi penyakit diabetes mellitus di Kecamatan Buleleng?

## 1.3 Tujuan Penelitian

Sesuai latar belakang yang telah dipaparkan, adapun tujuan dari kajian ini sebagai berikut.

1. Untuk mendeskripsikan hasil penerapan algoritma *K-Nearest Neighbor* dalam mengklasifikasi penyakit diabetes mellitus di Kecamatan Buleleng.

2. Untuk merancang penerapan algoritma *K-Nearest Neighbor* dalam mengklasifikasi penyakit diabetes mellitus di Kecamatan Buleleng.
3. Untuk mengimplementasikan penerapan algoritma *K-Nearest Neighbor* dalam mengklasifikasi penyakit diabetes mellitus di Kecamatan Buleleng.
4. Untuk mengetahui keakurasian aplikasi yang menerapkan algoritma *K-Nearest Neighbor* dalam mengklasifikasi penyakit diabetes mellitus di Kecamatan Buleleng.

#### **1.4 Manfaat Hasil Penelitian**

Melalui adanya kajian ini sehingga nantinya bisa memberi manfaat-manfaat berikut.

##### **1.4.1 Manfaat Teoretis**

Adapun manfaat teoritis yang diinginkan oleh peneliti dalam melakukan kajian ini yaitu bisa memberi ide melalui pemikiran serta dapat menambah kasanah ilmu pengetahuan dalam bidang matematika khususnya mengenai metode klasifikasi menggunakan algoritma *K-Nearest Neighbor* pada identifikasi penyakit diabetes mellitus.

##### **1.4.2 Manfaat Praktis**

###### **a. Bagi Penulis**

Manfaat dari kajian ini untuk penulis yakni bisa memperkaya wawasan khususnya pada konteks materi *data mining* mengenai metode klasifikasi untuk melakukan pengidentifikasian penyakit diabetes mellitus mempergunakan algoritma *K-Nearest Neighbor*.

b. Bagi Pembaca

Kajian ini bisa memperkaya pengetahuan untuk pembaca serta mampu dimanfaatkan selaku referensi ketika melakukan kajian berikutnya.

c. Bagi Pemerintah

Hasil dari kajian ini diharapkan bisa memberikan manfaat kepada pihak-pihak instansi kesehatan, sehingga nantinya kajian ini bisa dipergunakan selaku suatu acuan untuk pengambilan suatu keputusan dari proses metode klasifikasi pada identifikasi terkait penyakit diabetes mellitus.

### 1.4.3 Keterbatasan Masalah

Supaya problematika yang ada lebih terarah serta tidak melenceng dari pokok masalah, maka diperlukan batasan permasalahan, meliputi.

1. Parameter yang dipergunakan pada kajian ini meliputi jenis kelamin, riwayat penyakit tidak menular dalam diri sendiri, riwayat penyakit tidak menular dalam keluarga, berat badan, tinggi badan, *body mass index* (BMI), usia, lingkar perut, pemeriksaan gula, dan diagnosis.
2. Melakukan pengolahan data menggunakan format file CSV (*Comma-Separated Values*), yang digunakan untuk menyimpan data dalam bentuk tabular. Setiap baris dalam file CSV akan mewakili satu baris dalam tabel, dan setiap kolom akan dipisahkan oleh tanda koma.

3. Keterbatasan dalam deployment: Model aplikasi prediksi penyakit diabetes yang dibangun menggunakan model *waterfall* saat ini belum melalui tahapan *deployment*. Oleh karena itu, aplikasi yang telah dibuat hanya dapat digunakan pada PC dengan mengirimkan file *source code* aplikasi secara manual. Hal ini mengakibatkan ketergantungan pada lingkungan pengembangan karena model aplikasi yang hanya dapat diakses dan digunakan dalam lingkungan pengembangan yang spesifik menghadirkan keterbatasan dalam hal aksesibilitas.

