

BAB I PENDAHULUAN

1.1. Latar Belakang

Diabetes melitus merupakan penyakit gangguan metabolisme yang ditandai dengan kenaikan kadar gula dalam darah. Penyakit ini tergolong berbahaya, karena dalam jangka waktu lama dapat menyebabkan kerusakan ginjal, jaringan saraf, disfungsi mata, dan pembuluh darah (Hardianto, n.d.). Menurut data *The International Diabetes Federation*, diabetes termasuk salah satu masalah kesehatan dengan pertumbuhan paling cepat di abad ke-21. Di Indonesia prevalensi diabetes pada usia antara 20-79 tahun sekitar 10,6% yang menandakan 1 dari 9 orang terkena diabetes (M. Ratna Saraswati & I. Ngoerah, 2022).

Di Pos Pembinaan Terpadu, Skrining Penyakit Tidak Menular (Posbindu) Desa Bulupitu pada kurun waktu Januari hingga Februari 2024, mencatat lebih dari 80 pasien mengidap diabetes. Sebagian dari pasien tersebut mengetahui dirinya terkena diabetes dalam kondisi yang lumayan parah. Hal ini karena kurangnya kesadaran masyarakat untuk *check-up* kesehatan secara berkala, sedangkan pola hidup yang dijalani terbilang tidak sehat. Di Posbindu Desa Bulupitu masyarakat hanya bisa mengecek gula darah. Sedangkan untuk mendeteksi diabetes, perlu pemeriksaan hemoglobin A1c (*HbA1c*). Pemeriksaan ini memiliki kaitan terhadap kadar glukosa dalam darah pada penderita diabetes melitus (Hartini, 2016). Namun, layanan pengecekan *HbA1c* belum tersedia di Posbindu Desa Bulupitu, Kabupaten Malang.

Posbindu Desa Bulupitu memerlukan pendekatan lain untuk mendeteksi dini resiko Masyarakat terkena diabetes yakni dengan data mining. Metode data mining terdiri oleh serangkaian alat dan 1 teknik yang mampu mengeksplorasi kumpulan data dan membantu penemuan pengetahuan (L. Barreto Moreira & A. Amendoeira Namen, 2018). Beberapa metode yang bisa digunakan untuk memprediksi kemungkinan diabetes antara lain: *random forest* [5](Aprilia et al., 2021), *support vector machine* (A. Dwi Cahyani & A. Basuki, 2019), *K-Nearest Neighbor (KNN)* (Dwi Fasnuari et al., 2022), dan *Naïve Bayes* (C. A. Rahayu et al., 2023). Dalam penelitian ini penulis akan membandingkan metode *KNN* dan *Naïve Bayes* untuk memprediksi diabetes di Posbindu Desa Bulupitu.

Pada penelitian (Dwi Fasnuari et al., 2022), peneliti menggunakan *KNN* untuk klasifikasi diabetes. Perhitungan jarak yang digunakan hanya *Euclidean distance*. Pada penelitian ini penulis akan membandingkan beberapa perhitungan jarak, antara lain *Euclidean distance*, *Mahanttan distance*, dan *Chebyshev distance*.

1.2. Rumusan Masalah

1. Bagaimana performa metode *Naïve Bayes* dalam memprediksi risiko diabetes di Posbindu Desa Bulupitu?
2. Bagaimana performa metode *K-Nearest Neighbor (KNN)* dalam memprediksi risiko diabetes di Posbindu Desa Bulupitu?
3. Manakah dari kedua metode tersebut yang memiliki akurasi lebih tinggi dalam memprediksi risiko diabetes?

1.3. Tujuan

Tujuan utama penelitian ini adalah membandingkan model prediksi menggunakan *Naïve Bayes* dan *K-Nearest Neighbor (KNN)* untuk mengidentifikasi pasien yang berisiko mengidap diabetes saat berkunjung ke Posbindu Desa Bulupitu. Secara khusus, kami bertujuan untuk:

1. Menganalisis performa metode *Naïve Bayes* dalam memprediksi risiko diabetes di Posbindu Desa Bulupitu.
2. Menganalisis performa metode *K-Nearest Neighbor (KNN)* dalam memprediksi risiko diabetes di Posbindu Desa Bulupitu.
3. Membandingkan akurasi kedua metode untuk menentukan metode yang lebih efektif dalam memprediksi risiko diabetes.

1.4. Manfaat Penelitian

1. Manfaat Teoritis: Memberikan kontribusi pada pengembangan ilmu pengetahuan di bidang data mining, khususnya dalam klasifikasi risiko diabetes menggunakan metode *Naïve Bayes* dan *KNN*.
2. Manfaat Praktis: Memberikan alternatif metode prediksi yang dapat digunakan di Posbindu Desa Bulupitu untuk mendeteksi risiko diabetes secara dini.

3. Manfaat Sosial: Meningkatkan kesadaran masyarakat Desa Bulupitu terhadap pentingnya pemeriksaan kesehatan secara berkala dan membantu mencegah komplikasi diabetes melalui deteksi dini.

1.5. Batasan Penelitian

- Penelitian ini hanya dilakukan di Posbindu Desa Bulupitu, Kabupaten Malang.
- Data yang digunakan terbatas pada data pasien yang berjumlah 100 orang.
- Variabel yang digunakan dalam penelitian ini meliputi jenis kelamin, usia, keturunan diabetes, sering kencing, *Body Mass Index (BMI)*, dan kadar gula darah.
- Pembagian dataset dilakukan menggunakan *cross-validation* dan beberapa model *rasio* (60:40, 70:30, 80:20, dan 90:10).