

BAB I PENDAHULUAN

1.1 Latar Belakang

Diabetes adalah kondisi kronis yang ditandai dengan tingginya kadar gula darah akibat gangguan dalam produksi atau fungsi insulin. Kondisi ini dapat menyebabkan serangkaian komplikasi serius seperti penyakit jantung, gangguan ginjal, dan kerusakan saraf [1]. Pada tahun 2021 penderita diabetes mencapai angka yang cukup tinggi yaitu 19,5 juta jiwa dengan prediksi pada tahun 2045 akan menyentuh angka 28,6 juta jiwa [2]. Menurut *World Health Organization* (WHO) meskipun risiko terkena penyakit diabetes yang tinggi, setengah dari penderita diabetes tidak menyadarinya. Maka dibutuhkan alternatif lain yang memudahkan dalam mendeteksi penyakit diabetes sehingga dapat memberikan peringatan dini, dan masyarakat bisa segera memperbaiki pola hidup agar terhindar dari penyakit diabetes.

Terdapat berbagai jenis algoritma yang bisa digunakan untuk mengidentifikasi dan membuat *machine learning*. Salah satunya algoritma yang digunakan untuk mengidentifikasi jenis penyakit diabetes adalah algoritma *random forest*. Metode *random forest* merupakan bagian dari metode data mining dengan metode dasarnya pohon keputusan (*decision tree*). *Random forest* membangun suatu model dengan menggunakan beberapa pohon keputusan (*decision tree*) secara acak dan menggabungkan prediksi setiap pohon untuk mendapatkan hasil prediksi [3]. Kelebihan algoritma *Random Forest* yaitu mampu memberikan akurasi yang sangat tinggi meskipun terdapat data yang hilang serta sangat efisien untuk menyimpan data, *Random Forest* juga memiliki fitur seleksi yang dimana mampu untuk mengambil fitur terbaik dan menghasilkan hasil terbaik terhadap model klasifikasi [4].

Dataset yang tidak seimbang biasanya akan mempengaruhi hasil. Maka digunakan *Synthetic Minority Oversampling Technique* (SMOTE) untuk menyeimbangkan data tersebut. Pada SMOTE, data pada kelas minoritas diperbanyak dengan menciptakan data sintetis yang didasarkan pada replikasi data

dari kelas minoritas [5].

Berdasarkan latar belakang yang telah diuraikan di atas, maka peneliti akan melakukan penelitian dengan judul “**Deteksi Penyakit Diabetes Menggunakan Synthetic Minority Oversampling Technique Dan Metode Random Forest**”.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka pada penelitian ini dirumuskan masalah sebagai berikut:

- Bagaimana metode *random forest* bisa diterapkan dalam mendeteksi penyakit diabetes?
- Berapa tingkat akurasi yang diperoleh dari algoritma *random forest* dengan SMOTE untuk mendeteksi penyakit diabetes?
- Bagaimana pengimplementasiannya menjadi sebuah aplikasi website?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini meliputi:

- Menggunakan metode *random forest* klasifikasi.
- Dataset yang digunakan merupakan dataset penyakit diabetes yang diambil dari Kaggle dengan sumber <https://www.kaggle.com/datasets/ehababoelnaga/diabetes-dataset>.
- Menggunakan SMOTE untuk menyeimbangkan kelas dari dataset.
- Google collab digunakan untuk modeling dan Bahasa pemrograman yang digunakan adalah PYTHON.
- Visual studio code untuk pengembangan website dengan menggunakan bahasa pemrograman HTML dan CSS.
- Menggunakan *framework* Streamlite untuk mengubah skrip data menjadi aplikasi website.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh peneliti dalam penelitian ini adalah website pendeteksi penyakit diabetes menggunakan metode random forest.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah untuk memberikan efisiensi untuk tenaga medis dalam memberikan diagnosis awal penyakit diabetes secara mudah, hanya cukup dengan mengaksesnya melalui sebuah website.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Berisi Latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada Bab ini berisi tinjauan studi literatur, dasar teori.

BAB III METODE PENELITIAN

Pada Bab III alur penelitian, alat dan bahan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini merupakan tahapan yang penulis lakukan dalam mengembangkan aplikasi, testing hingga penerapan aplikasi di pendeteksi penyakit diabetes.

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran untuk peneliti selanjutnya.