

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Diabetes Melitus (DM) adalah kelompok gangguan metabolisme yang ditandai dengan tingginya kadar glukosa dalam darah (*hiperglikemia*). Kondisi ini disebabkan oleh gangguan dalam sekresi insulin, fungsi insulin, atau keduanya. Hiperglikemia kronis pada diabetes berhubungan dengan komplikasi jangka panjang yang dapat merusak atau menyebabkan kegagalan pada berbagai organ tubuh, terutama mata, ginjal, saraf, jantung dan pembuluh darah (Scobie, 2019). Menurut laporan dari Kementerian Kesehatan Republik Indonesia (2020), jumlah kasus Diabetes Melitus di Indonesia mencapai 10,9% dari total populasi, dengan mayoritas kasus terdiagnosis pada tahap lanjut. Penyakit ini dapat menyebabkan komplikasi serius seperti penyakit jantung, stroke, dan gagal ginjal, yang menjadi salah satu faktor peningkatan angka kematian serta penurunan kualitas hidup penderitanya.

Salah satu tantangan besar dalam menangani Diabetes Melitus adalah kurangnya kesadaran masyarakat untuk melakukan pemeriksaan kesehatan secara rutin, terutama dalam memprediksi risiko diabetes sejak dini. Banyak penderita diabetes yang baru menyadari kondisinya ketika komplikasi sudah muncul, mengakibatkan penanganan yang terlambat dan sulit dilakukan. Menurut *American Diabetes Association* (2021), pencegahan dan deteksi dini risiko diabetes memainkan peran penting dalam mengurangi beban penyakit ini, terutama melalui identifikasi faktor risiko seperti indeks massa tubuh (BMI), tekanan darah, dan gaya hidup.

Dengan perkembangan teknologi di bidang kesehatan, khususnya pemanfaatan *machine learning*, muncul peluang untuk membantu masyarakat melakukan deteksi dini risiko Diabetes Melitus dengan lebih mudah dan efisien. *Machine learning* adalah cabang dari kecerdasan buatan (*artificial intelligence*) yang memungkinkan sistem untuk belajar dari data dan membuat prediksi atau keputusan tanpa harus diprogram secara eksplisit. Penelitian oleh Verma & Verma (2022) menjelaskan bahwa *machine learning* dapat digunakan untuk meningkatkan

efektivitas dalam pengambilan keputusan medis dengan analisis data yang lebih baik dan lebih cepat.

Salah satu tantangan utama dalam penerapan *machine learning* untuk prediksi risiko Diabetes Melitus adalah ketidakseimbangan kelas (*class imbalance*) pada dataset. Ketidakseimbangan ini terjadi ketika jumlah data kelas mayoritas (*non-diabetes*) jauh lebih dominan dibandingkan kelas minoritas (*prediabetes* atau *diabetes*), sehingga model cenderung lebih mudah mengenali kelas mayoritas dan mengalami kesulitan dalam mengidentifikasi kelas minoritas (Abushahla & Pala, 2024). Kondisi tersebut dapat memengaruhi performa model, khususnya dalam mendeteksi individu yang berada pada kelompok berisiko. Oleh karena itu, diperlukan penanganan *class imbalance* agar model dapat mempelajari karakteristik setiap kelas secara lebih seimbang.

Dalam penelitian ini, algoritma *Random Forest* (RF) dipilih sebagai model prediksi utama karena memiliki kemampuan yang baik dalam menangani data dengan jumlah atribut yang banyak, hubungan antar variabel yang kompleks, serta mampu menghasilkan performa klasifikasi yang stabil. Selain itu, *Random Forest* juga dapat dikombinasikan dengan teknik penanganan *class imbalance* untuk meningkatkan kemampuan model dalam mengenali kelas minoritas. Penelitian oleh Chowdhury dkk., (2024) menunjukkan bahwa penerapan teknik penanganan *class imbalance* pada dataset diabetes BRFSS 2015 mampu meningkatkan performa model dalam mengklasifikasikan setiap kelas. Oleh karena itu, penelitian ini tidak hanya mengimplementasikan algoritma *Random Forest* pada aplikasi berbasis website, tetapi juga menganalisis kinerja model *multiclass* serta pengaruh penanganan *class imbalance* terhadap performa model.

Penelitian oleh König & Suhr, (2023) menunjukkan bahwa penggunaan aplikasi berbasis website memiliki potensi besar dalam meningkatkan literasi kesehatan digital dan kesadaran masyarakat terhadap pentingnya pencegahan penyakit. Website sebagai platform memberikan akses yang luas dan fleksibilitas bagi pengguna untuk memperoleh informasi kesehatan secara mandiri.

Dengan demikian, penelitian ini berkontribusi pada dua aspek utama, yaitu pengembangan aplikasi berbasis website untuk deteksi dini risiko diabetes serta analisis kinerja algoritma *Random Forest* pada klasifikasi *multiclass* dengan

mempertimbangkan pengaruh penanganan *class imbalance* terhadap performa model.

1.2 Perumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

- 1) Bagaimana mengimplementasikan algoritma *Random Forest* pada aplikasi website untuk mendeteksi risiko diabetes?
- 2) Bagaimana kinerja algoritma *Random Forest* dalam mendeteksi risiko diabetes pada klasifikasi *multiclass* (non-diabetes, prediabetes, dan diabetes) ?
- 3) Bagaimana pengaruh penanganan *class imbalance* terhadap performa algoritma *Random Forest* dalam mendeteksi risiko diabetes?

1.3 Batasan Masalah

Agar ruang lingkup penelitian tidak terlalu luas, batasan masalah yang diterapkan adalah sebagai berikut:

- 1) Aplikasi ini hanya memprediksi risiko Diabetes Melitus dan tidak memberikan diagnosis medis resmi.
- 2) Data yang digunakan untuk prediksi didasarkan pada *input* pengguna terkait faktor kesehatan, seperti *BMI*, tekanan darah dan aktivitas fisik.
- 3) Pengujian aplikasi terbatas pada dataset kesehatan yang tersedia serta uji coba fungsionalitas pada platform *website*.
- 4) Implementasi algoritma hanya dilakukan pada platform *website*.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Penelitian ini bertujuan untuk:

- 1) Mengembangkan aplikasi berbasis website yang mampu melakukan prediksi risiko Diabetes Melitus menggunakan algoritma *Random Forest*.
- 2) Membangun model klasifikasi *multiclass* untuk mengelompokkan tingkat risiko Diabetes Melitus berdasarkan data faktor-faktor risiko yang digunakan dalam penelitian.
- 3) Menerapkan teknik penanganan *class imbalance* pada dataset guna meningkatkan kemampuan model dalam mengenali kelas minoritas.

- 4) Menganalisis pengaruh penerapan teknik *class imbalance* terhadap performa algoritma *Random Forest* pada klasifikasi *multiclass*.
- 5) Mengevaluasi performa model berdasarkan matriks evaluasi seperti *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*.
- 6) Menyediakan media yang dapat membantu masyarakat melakukan deteksi dini risiko Diabetes Melitus melalui aplikasi berbasis website.

1.4.2 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat bagi pengembangan ilmu pengetahuan maupun penerapan teknologi di bidang prediksi risiko Diabetes Melitus. Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

- 1) Menambah pengetahuan dan pengalaman dalam menerapkan algoritma *Random Forest* serta teknik *class imbalance* untuk prediksi risiko Diabetes Melitus.
- 2) Menjadi referensi bagi mahasiswa dan peneliti yang ingin mengembangkan penelitian mengenai prediksi risiko Diabetes Melitus menggunakan metode *machine learning*.
- 3) Menjadi bahan perbandingan dalam penelitian yang membahas penerapan teknik *class imbalance* pada permasalahan klasifikasi data kesehatan.
- 4) Menghasilkan aplikasi berbasis website yang dapat digunakan untuk melakukan prediksi risiko Diabetes Melitus berdasarkan data yang dimasukkan oleh pengguna.
- 5) Menunjukkan implementasi algoritma *Random Forest* pada sistem prediksi berbasis website yang dapat dikembangkan lebih lanjut pada penelitian berikutnya.