

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Kesehatan ibu hamil merupakan salah satu prioritas utama dalam sistem kesehatan global karena berperan langsung terhadap kualitas generasi mendatang. Salah satu permasalahan yang semakin mendapat perhatian adalah meningkatnya kasus Diabetes Mellitus Gestasional (DMG), yaitu kondisi gangguan metabolisme glukosa yang terjadi selama masa kehamilan. Organisasi kesehatan dunia melaporkan bahwa prevalensi diabetes gestasional terus mengalami peningkatan seiring perubahan gaya hidup, pola konsumsi, serta meningkatnya usia kehamilan pada perempuan (Adli dkk., 2021).

Secara empiris, diperkirakan sekitar 14% kehamilan di seluruh dunia terdampak oleh diabetes gestasional, atau setara dengan sekitar 18 juta kelahiran setiap tahunnya (Plows dkk., 2018). Angka ini berkaitan erat dengan meningkatnya prevalensi obesitas, gaya hidup sedentari, dan pola makan yang tidak sehat pada perempuan usia subur (Plows dkk., 2018). Diabetes gestasional menjadi perhatian serius karena tidak hanya berdampak pada kesehatan ibu, tetapi juga meningkatkan risiko komplikasi pada janin, seperti makrosomia, kelahiran prematur, hingga gangguan metabolik jangka panjang pada anak (Adli dkk., 2021). Oleh karena itu, diperlukan upaya deteksi dini yang lebih efektif guna meminimalkan risiko komplikasi yang dapat terjadi.

Ibu hamil dipilih sebagai objek penelitian karena merupakan kelompok yang memiliki kerentanan khusus terhadap diabetes gestasional. Perubahan hormonal selama kehamilan dapat menyebabkan resistensi kadar insulin, sehingga meningkatkan risiko terjadinya gangguan metabolisme glukosa meskipun sebelumnya tidak memiliki riwayat diabetes. Selain itu, kondisi diabetes gestasional bersifat unik karena hanya terjadi pada masa kehamilan, namun memiliki dampak jangka panjang baik bagi ibu maupun anak yang dilahirkan (Plows dkk., 2018).

Pada kondisi ideal, setiap ibu hamil seharusnya menjalani skrining secara rutin untuk mendeteksi risiko diabetes gestasional sejak dini. Namun, pada praktiknya, proses diagnosis masih bergantung pada pemeriksaan laboratorium

seperti Tes Toleransi Glukosa Oral (TTGO) yang memerlukan waktu, biaya, serta ketersediaan fasilitas kesehatan yang memadai. Keterbatasan tersebut menyebabkan tidak semua ibu hamil mendapatkan pemeriksaan secara tepat waktu, sehingga deteksi risiko sering terlambat dilakukan (Adli dkk., 2021). Oleh karena itu, diperlukan suatu pendekatan alternatif yang mampu membantu proses skrining awal secara lebih cepat dan efisien.

Pada konteks pengembangan sistem berbasis Teknik Informatika, pemanfaatan data rekam medis melalui teknik data mining dan *machine learning* menjadi salah satu solusi yang potensial untuk membantu proses deteksi dini penyakit, termasuk diabetes (Kavakiotis dkk., 2017). *Dataset* medis yang berisi data klinis ibu hamil seperti jumlah kehamilan, kadar glukosa, tekanan darah, Indeks Massa Tubuh, riwayat keturunan diabetes, dan usia telah banyak digunakan untuk membangun model klasifikasi risiko diabetes secara komputasional (Delvika dkk., 2022).

Salah satu metode yang dapat digunakan adalah algoritma *Naive Bayes*, yang dikenal memiliki keunggulan dalam hal kesederhanaan, efisiensi, serta kemampuan dalam menangani data dengan jumlah terbatas (Sihananto & Maulana, 2021). Namun, dalam kasus data medis, sering ditemukan permasalahan ketidakseimbangan data (*imbalanced data*), di mana jumlah data pada satu kelas lebih dominan dibandingkan kelas lainnya. Kondisi ini dapat menyebabkan model menjadi bias terhadap kelas mayoritas dan kurang optimal dalam mengenali kelas minoritas (Suliswati dkk., 2025).

Mengatasi permasalahan tersebut, digunakan teknik *Random Over Sampling* (ROS) yang bertujuan untuk menyeimbangkan jumlah data antar kelas dengan cara menduplikasi data pada kelas minoritas. Penelitian terdahulu menunjukkan bahwa kombinasi ROS dengan *Naive Bayes* efektif meningkatkan sensitivitas dan spesifisitas model dalam mendeteksi kelas minoritas pada data kesehatan yang tidak seimbang (Suliswati dkk., 2025). Melalui penerapan ROS, diharapkan model *Naive Bayes* dapat bekerja lebih optimal dalam mendeteksi kelas minoritas, khususnya dalam mengidentifikasi ibu hamil yang berisiko mengalami diabetes gestasional.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk membangun model klasifikasi risiko diabetes mellitus gestasional pada ibu hamil menggunakan

algoritma *Naive Bayes* yang dikombinasikan dengan teknik *Random Over Sampling* guna meningkatkan akurasi dan kemampuan deteksi model.

1.2 Rumusan Penelitian

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian dijabarkan sebagai berikut.

1. Bagaimana mengimplementasikan algoritma *Naive Bayes Classifier* untuk mengklasifikasikan risiko diabetes gestasional pada ibu hamil?
2. Seberapa besar tingkat akurasi, presisi, dan *recall* yang dihasilkan oleh algoritma *Naive Bayes* dalam memprediksi risiko tersebut menggunakan dataset yang tersedia?

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini sesuai dengan rumusan masalah sebelumnya adalah sebagai berikut.

1. Membangun model klasifikasi cerdas untuk mengklasifikasikan risiko diabetes gestasional menggunakan algoritma *Naive Bayes*.
2. Mengukur kinerja model tersebut untuk memastikan keandalan hasil prediksi menggunakan pengujian *confusion matrix* untuk mencari nilai akurasi, presisi, *recall*

1.4 Batasan Penelitian

Penelitian ini memiliki Batasan guna membatasi penelitian ini yang diuraikan sebagai berikut.

1. Penelitian ini hanya menggunakan metode *Naive Bayes Classifier* untuk proses klasifikasi.
2. *Dataset* yang digunakan adalah *dataset* sekunder *Gestational Diabetes Mellitus* (GDM Data Set). *Dataset* yang diperoleh dari repository Kaggle data ini berisi 3.525 data dan data ini berasal dari warga wilayah India data ini dibuat oleh Sumathi Santhosh pada tahun 2022 tersedia

<https://www.kaggle.com/datasets/sumathisanthosh/gestational-diabetes-mellitus-gdm-data-set>.

3. Atribut yang dianalisis dibatasi pada fitur-fitur klinis utama yaitu: *Age*, *No of Pregnancy*, *Gestation in Previous Pregnancy*, BMI, HDL, *Family History*,

Unexplained Prenatal Loss, Large Child or Birth Defect, PCOS, Sys BP, Dia BP, OGTT, Hemoglobin, Sedentary Lifestyle, Prediabetes

4. *Output* dari sistem ini adalah klasifikasi risiko dalam dua kategori yaitu: Berisiko dan Tidak Berisiko.
5. Penelitian ini hanya menggunakan metode filter *Random Over Sampling* untuk menyeimbangkan data.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat di masa mendatang yang disebutkan sebagai berikut.

1. Bagi Penulis: Mengaplikasikan ilmu kecerdasan buatan dan pengolahan data untuk menyelesaikan permasalahan nyata di bidang kesehatan.
2. Bagi Tenaga Kesehatan: Memberikan alat bantu tambahan dalam proses skrining awal sehingga penanganan pasien bisa dilakukan lebih efisien.
3. Bagi Akademisi: Memberikan kontribusi berupa literatur penelitian mengenai efektivitas algoritma *Naive Bayes* pada kasus medis spesifik wanita.

