

Pendekatan *Naïve Bayes* Untuk Klasifikasi Data Pasien Diabetes Gestasional

Fikri Ubaidillah^{1*}, Zezen Mujani², Nurul Amin³, M. Fikry Akbari⁴

¹⁻⁴Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspipetek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: 1*fikriubaidillah73@gmail.com

(* : coressponding author)

Abstrak- Diabetes gestasional adalah kondisi diabetes yang berkembang selama kehamilan dan dapat menyebabkan komplikasi serius bagi ibu maupun janin jika tidak terdeteksi dan ditangani secara tepat. Identifikasi dini risiko diabetes gestasional menjadi penting untuk mendukung pencegahan dan penanganan yang lebih efektif. Dalam penelitian ini, metode Naive Bayes digunakan untuk mengklasifikasikan data pasien diabetes gestasional berdasarkan variabel-variabel kesehatan dan demografi seperti usia, indeks massa tubuh (BMI), riwayat keluarga diabetes, kadar gula darah, tekanan darah, dan hasil pemeriksaan laboratorium lainnya. Metode Naive Bayes dipilih karena kemampuannya dalam menangani dataset yang kompleks dan memberikan hasil klasifikasi dengan akurasi yang cukup tinggi, meskipun asumsinya yang sederhana. penelitian ini menunjukkan bahwa pendekatan Naive Bayes merupakan metode yang efektif untuk klasifikasi risiko diabetes gestasional, memberikan hasil yang stabil dan akurat. Hal ini menjadikannya alat yang berpotensi digunakan oleh tenaga medis sebagai pendukung keputusan dalam identifikasi awal risiko diabetes pada ibu hamil, yang selanjutnya dapat membantu dalam merancang intervensi yang sesuai untuk mencegah komplikasi.

Kata Kunci: *Naive Bayes, Diabetes Gestasional, Klasifikasi, Probabilitas, Machine Learning*

Abstract- *Gestational diabetes is a diabetes condition that develops during pregnancy and can cause serious complications for the mother and fetus if not detected and treated appropriately. Early identification of the risk of gestational diabetes is important to support more effective prevention and treatment. In this study, the Naive Bayes method was used to classify gestational diabetes patient data based on health and demographic variables such as age, body mass index (BMI), family history of diabetes, blood sugar levels, blood pressure, and the results of other laboratory tests. The Naive Bayes method was chosen because of its ability to handle complex datasets and provide classification results with fairly high accuracy, despite its simple assumptions. This study shows that the Naive Bayes approach is an effective method for gestational diabetes risk classification, providing stable and accurate results. This makes it a tool that can potentially be used by medical personnel as decision support in the early identification of diabetes risk in pregnant women, which can then help in designing appropriate interventions to prevent complications.*

Keywords: *Naive Bayes, Gestational Diabetes, Classification, Probability, Machine Learning*

1. PENDAHULUAN

Diabetes gestasional merupakan jenis diabetes yang muncul selama masa kehamilan dan dapat menimbulkan berbagai komplikasi serius bagi ibu serta janin, termasuk risiko preeklampsia, kelahiran prematur, dan komplikasi metabolik pada bayi baru lahir. Menurut Organisasi Kesehatan Dunia (WHO), prevalensi diabetes gestasional mengalami peningkatan di berbagai negara, terutama dengan meningkatnya jumlah kehamilan pada wanita yang memiliki risiko tinggi seperti obesitas atau riwayat keluarga dengan diabetes. Dalam upaya untuk mengurangi risiko ini, penting untuk melakukan deteksi dini terhadap ibu hamil yang memiliki potensi mengembangkan diabetes gestasional agar dapat segera diberikan intervensi yang sesuai (Ibrahim, 2017).

Salah satu cara untuk melakukan deteksi dini adalah dengan menggunakan metode klasifikasi berbasis machine learning, yang memungkinkan pemrosesan data pasien dalam jumlah besar dengan hasil prediktif yang dapat diandalkan.

Metode Naive Bayes adalah salah satu metode klasifikasi yang populer karena kesederhanaannya, kemampuannya dalam menangani data besar, dan efisiensi komputasinya. Naive Bayes bekerja dengan menggunakan prinsip probabilitas, di mana metode ini menghitung probabilitas suatu kelas berdasarkan distribusi fitur yang diasumsikan independen satu sama lain. Dalam konteks klasifikasi data medis, metode ini telah banyak diterapkan untuk memprediksi kondisi kesehatan tertentu dengan tingkat akurasi yang memadai, meskipun bekerja dengan asumsi yang sederhana.

Penelitian ini bertujuan untuk mengevaluasi efektivitas metode *Naive Bayes* dalam mengklasifikasikan data pasien yang memiliki risiko diabetes gestasional. Data yang digunakan meliputi variabel-variabel seperti usia, indeks massa tubuh (BMI), riwayat keluarga, tekanan darah, kadar gula darah, dan hasil pemeriksaan laboratorium lainnya. Dengan menggunakan metode *Naive Bayes*, diharapkan dapat dikembangkan model yang mampu memberikan prediksi risiko dengan akurasi tinggi dan waktu komputasi yang efisien, sehingga dapat diterapkan dalam lingkungan klinis untuk mendukung tenaga medis dalam pengambilan keputusan yang cepat dan tepat.

Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan model prediksi diabetes gestasional yang efektif, serta menjadi referensi bagi penelitian lebih lanjut dalam penerapan machine learning untuk bidang kesehatan.

2. METODE PENELITIAN

2.1 *Naive Bayes*

Naive Bayes bekerja dengan menggunakan probabilitas, di mana ia menghitung kemungkinan suatu data masuk ke dalam kelas tertentu berdasarkan fitur-fitur yang ada. Algoritma ini disebut “naive” karena mengasumsikan bahwa semua fitur atau variabel saling independen (Dwi) Artinya, algoritma menganggap bahwa keberadaan atau nilai dari satu fitur tidak terkait dengan fitur lainnya dalam menentukan hasil klasifikasi, yang mungkin tidak selalu benar dalam kenyataan. *Naive Bayes* memiliki kelebihan yang signifikan dalam mendukung diagnosis diabetes gestasional, terutama karena algoritma ini dapat dengan cepat memproses data medis untuk mengidentifikasi pola yang terkait dengan faktor risiko seperti usia, indeks massa tubuh (BMI), riwayat keluarga, kadar glukosa, dan tekanan darah.

Algoritma ini memungkinkan analisis probabilistik yang cepat dan sederhana, sehingga sangat bermanfaat untuk skrining awal dan diagnosis dini. Kecepatan dan efisiensinya sangat membantu dalam lingkungan klinis, di mana waktu respons yang cepat diperlukan untuk meminimalkan risiko komplikasi bagi ibu dan janin. Dengan asumsi independensi antar faktor risiko, *Naive Bayes* tetap mampu memberikan hasil klasifikasi yang akurat, menjadikannya alat bantu yang andal dalam menentukan pasien yang berisiko tinggi dan memerlukan penanganan lebih lanjut.

3. ANALISA DAN PEMBAHASAN

3.1 Diabetes Gestasional

Diabetes gestasional (gestational diabetes mellitus, GDM) adalah kondisi diabetes yang muncul pada wanita selama masa kehamilan dan biasanya menghilang setelah persalinan. Kondisi ini terjadi ketika tubuh tidak mampu memproduksi cukup insulin untuk mengatur kadar gula darah selama kehamilan, karena adanya perubahan hormon dan kebutuhan metabolik yang meningkat (Umiyah, 2023).

3.1.1 Penyebab Diabetes Gestasional

Metode yang digunakan pada pengumpulan data dalam program aplikasi ini adalah sebagai berikut: Diabetes gestasional disebabkan oleh perubahan hormon yang menghambat kerja insulin. Selama kehamilan, tubuh menghasilkan hormon yang mendukung perkembangan janin, seperti estrogen, kortisol, dan laktogen plasenta manusia, yang dapat mengganggu kerja insulin. Biasanya, pankreas merespons peningkatan kebutuhan insulin ini dengan meningkatkan produksinya, namun pada beberapa wanita, tubuh tidak bisa memenuhi kebutuhan insulin yang meningkat ini, sehingga terjadi peningkatan kadar gula darah (Maulana, 2023).

3.1.2 Faktor Risiko

Beberapa faktor risiko diabetes gestasional meliputi:

- Usia, Wanita berusia lebih dari 25 tahun memiliki risiko lebih tinggi.
- Riwayat Keluarga, Adanya riwayat diabetes pada keluarga.
- Obesitas, Wanita dengan indeks massa tubuh (IMT) tinggi memiliki risiko lebih besar.

- d. Kehamilan Sebelumnya, Pernah mengalami diabetes gestasional atau melahirkan bayi dengan berat lahir lebih dari 4 kg.
- e. Riwayat Sindrom Metabolik, Wanita yang memiliki sindrom metabolik atau riwayat tekanan darah tinggi.

3.2 Dampak pada Ibu dan Janin

Diabetes gestasional dapat berdampak pada ibu dan janin jika tidak dikelola dengan baik. Pada ibu, kondisi ini meningkatkan risiko preeklampsia, operasi caesar, dan diabetes tipe 2 di masa depan (Aprianto, 2023). Pada janin, diabetes gestasional dapat menyebabkan:

1. Makrosomia, Bayi dengan berat lahir yang sangat besar, yang dapat mempersulit proses persalinan.
2. Hipoglikemia Setelah lahir, bayi mungkin mengalami gula darah rendah.
3. Risiko Diabetes, Bayi yang lahir dari ibu dengan diabetes gestasional lebih mungkin mengembangkan obesitas dan diabetes tipe 2 di masa dewasa.

3.3 Penanganan Diabetes Gestasional

Manajemen diabetes gestasional meliputi pengaturan pola makan, olahraga, dan pemantauan kadar gula darah. Pada beberapa kasus, ibu mungkin perlu mengonsumsi obat antidiabetes atau suntikan insulin jika kadar gula darah tetap tinggi meskipun telah mengubah pola makan dan gaya hidup (Sulaiman, 2024).

3.5 Metode

Metode Naive Bayes adalah algoritma pembelajaran mesin berbasis probabilitas yang cocok untuk klasifikasi data dengan asumsi independensi antar fitur (Yunita, 2018). Naive Bayes sangat berguna dalam menganalisis data kesehatan, termasuk untuk klasifikasi risiko diabetes gestasional (GDM). Berikut adalah keterkaitan dan penerapan metode Naive Bayes pada data diabetes gestasional (Maulidina, 2024).

1. Tujuan Klasifikasi Risiko

Naive Bayes dapat digunakan untuk memprediksi apakah seorang wanita memiliki risiko tinggi atau rendah terkena diabetes gestasional berdasarkan sejumlah faktor risiko.

Beberapa variabel yang umumnya digunakan dalam analisis risiko diabetes gestasional meliputi:

- a. Usia Ibu, Wanita yang lebih tua cenderung memiliki risiko lebih tinggi.
- b. IMT (Indeks Massa Tubuh), Obesitas berkorelasi dengan risiko GDM yang lebih tinggi.
- c. Riwayat Keluarga, Adanya riwayat diabetes dalam keluarga dapat meningkatkan risiko.
- d. Riwayat GDM Sebelumnya, Wanita yang pernah mengalami diabetes gestasional sebelumnya lebih berisiko.
- e. Tekanan Darah Tinggi Adanya hipertensi atau sindrom metabolik.

2. Bagaimana Naive Bayes Menganalisis Risiko GDM

Metode Naive Bayes menghitung probabilitas dari setiap kelas (misalnya, kelas "berisiko" atau "tidak berisiko" untuk GDM) berdasarkan nilai atribut atau fitur yang diberikan (seperti usia, IMT, riwayat keluarga, dll.).

Dalam konteks GDM, Naive Bayes menghitung:

$$P(GDM|faktor_risiko) = \frac{P(faktor_risiko|GDM) \times P(GDM)}{P(faktor_risiko)}$$

Dimana:

- $P(GDM|faktor_risiko)$ adalah probabilitas seorang wanita mengalami diabetes gestasional dengan adanya faktor risiko tertentu.
- $P(faktor_risiko|GDM)$ adalah probabilitas memiliki faktor risiko tertentu jika wanita tersebut mengidap GDM.
- $P(GDM)$ adalah probabilitas dasar dari seorang mengalami GDM dalam Populasi.
- $P(faktor_risiko)$ adalah probabilitas memiliki faktor risiko tertentu tanpa mempertimbangkan GDM.

3. Kelebihan Naive Bayes untuk Data Kesehatan

- a. Efisiensi, Naive Bayes relatif cepat dan bekerja baik dengan data berukuran besar.
- b. Kemampuan Generalisasi, Meski fitur tidak benar-benar independen, Naive Bayes tetap menunjukkan performa yang baik, yang membantu dalam analisis data kesehatan yang cenderung memiliki banyak variabel korelasi.
- c. Interpretabilitas, Naive Bayes menghasilkan probabilitas yang mudah dipahami untuk setiap kelas, yang dapat membantu dalam memberikan penjelasan sederhana tentang risiko GDM pada pasien.

4. KESIMPULAN

Naive Bayes adalah metode pembelajaran mesin berbasis probabilitas yang sering digunakan dalam klasifikasi, terutama ketika tujuannya adalah untuk memprediksi kategori dari data tertentu, seperti risiko penyakit. Dalam konteks diabetes gestasional (GDM), Naive Bayes dapat dimanfaatkan untuk memperkirakan apakah seorang wanita hamil memiliki risiko tinggi atau rendah mengalami GDM berdasarkan faktor-faktor risiko yang ada.

GDM biasanya berkaitan dengan sejumlah faktor seperti usia ibu, indeks massa tubuh (IMT), riwayat keluarga dengan diabetes, riwayat GDM pada kehamilan sebelumnya, dan kondisi tekanan darah. Metode Naive Bayes bekerja dengan menghitung probabilitas seseorang terkena GDM berdasarkan kombinasi dari faktor-faktor ini. Sebagai contoh, jika seorang wanita hamil memiliki usia lebih dari 30 tahun, IMT yang tinggi, dan riwayat keluarga dengan diabetes, Naive Bayes dapat mengkalkulasi peluang bahwa dia berisiko mengalami diabetes gestasional. Perhitungan ini berdasarkan distribusi probabilitas masing-masing faktor untuk setiap kelas risiko (berisiko atau tidak berisiko).

Kelebihan Naive Bayes adalah kemampuannya untuk memberikan hasil yang cepat dan efisien, meskipun pada kasus data medis, beberapa fitur mungkin saling berkorelasi, seperti usia dan IMT. Meskipun asumsi independensi antar fitur tidak selalu sepenuhnya akurat dalam data kesehatan, Naive Bayes sering kali masih menghasilkan prediksi yang cukup akurat dalam klasifikasi risiko. Ini menjadikannya pilihan yang menarik untuk digunakan dalam analisis kesehatan.

Dengan Naive Bayes, institusi medis dapat melakukan skrining awal pasien untuk risiko GDM dengan lebih cepat. Prediksi risiko ini bisa dijadikan acuan awal dalam menentukan perlunya pemeriksaan atau tindakan medis lebih lanjut. Oleh karena itu, Naive Bayes tidak hanya berguna untuk memperkirakan risiko diabetes gestasional tetapi juga mendukung dokter dalam proses pengambilan keputusan awal yang efektif dan tepat waktu.

Hubungan antara Naive Bayes dan diabetes gestasional adalah bahwa algoritma ini menawarkan pendekatan sederhana, cepat, dan terukur untuk memprediksi risiko GDM. Meskipun terbatas oleh asumsi independensi fitur, Naive Bayes tetap dapat memberikan estimasi yang berguna dan akurat dalam mendukung keputusan medis awal.

REFERENCES

- E. Moriesta and A. Ibrahim, (2017). "Prosiding Annual Research Seminar,".
- A. Dwi and P. Informatika, "SISTEM PAKAR DIAGNOSA PENYAKIT KUCING PERSIA MENGGUNAKAN METODE NAIVE BAYES."
- P. Sofyan Zakaria, R. Julianto, and R. Surya Bernada, "IMPLEMENTASI NAIVE BAYES MENGGUNAKAN PYTHON DALAM KLASIFIKASI DATA." [Online]. Available: <https://jurnalmahasiswa.com/index.php/biikma>
- A. Umiyah, "Analisis kejadian diabetes melitus gestasional di wilayah kerja Puskesmas Banyuputih," *Jurnal Ilmu Kesehatan Bhakti Husada: Health Sciences Journal*, vol. 14, no. 02, pp. 317–323, Dec. 2023, doi: 10.34305/jikbh.v14i02.824.
- R. Maulana, M. A. Fadillah, S. S. Pambudi, P. Rosyani, and F. I. Komputer, (2023). "Literature Review: Implementasi Sistem Pakar Untuk Diagnosa Penyakit Diabetes Menggunakan Metode Fuzzy," *Teknik dan Multimedia*, vol. 1, no. 2.

- F. Kamali Adli, C. Author, P. Pendidikan Dokter, F. Kedokteran, and U. Lampung, “*DIABETES MELITUS GESTASIONAL: DIAGNOSIS DAN FAKTOR RISIKO*.” [Online]. Available: <http://jurnalmedikahutama.com>
- N. Ningsih, A. Aprianto, and A. Angeline, (2023). “Pendekatan Data Science untuk Deteksi Dini Diabetes Menggunakan Naive Bayes Classifier,” *Journal of Information System, Graphics, Hospitality and Technology*, vol. 5, no. 1, pp. 26–31, doi: 10.37823/insight.v5i1.300.
- J. Khatib Sulaiman, A. Khoirala Ermy Pily, F. Aprilia, L. Efrizoni, and I. Artikel Abstrak, “Komparasi Algoritma K-Nearest Neighbors dan Naïve Bayes dalam Klasifikasi Penyakit Diabetes Gestasional STMIK Amik Riau,” *Indonesian Journal of Computer Science Attribution*, vol. 13, no. 1, pp. 2024– 1195.
- D. Yunita, R. Amalia, and D. P. Rosyani, (2018). “Analisa Prestasi Siswa Berdasarkan Kedisiplinan, Nilai Hasil Belajar, Sosial Ekonomi dan Aktivitas Organisasi Menggunakan Algoritma Naïve Bayes,” vol. 3, no. 4, pp. 2622–4615, [Online]. Available: <http://openjournal.unpam.ac.id/index.php/informatika>
- E. Job, N. Maulidina, M. Simanjuntak, and Y. Maulita, (2024). “THE USE OF BAYES METHOD TO DIAGNOSE GESTATIONAL DIABETES IN PREGNANT WOMEN (CASE STUDY: DR,” *Journal of Mathematics and Technology (MATECH)*, vol. 3, no. 2.