JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096

Pendekatan *Naïve Bayes* Untuk Klasifikasi Data Pasien Diabetes Gestasional

Perani Rosyani[1], Ilyas Saputra[2*], Niko Kristiawan[3], Abdus Salam[4], Harry Aprianto[5]

[1,2,3,4,5]Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia
Email: [1]dosen00837@unpam.ac.id, [2*]Ilyassaputra3210@gmail.com, [3]Nikstrom31@gmail.com, [4]Abduswae0102@gmail.com, [5]Harryapriantoaa71@gmail.com
(* : coressponding author)

Abstrak - Diabetes gestasional merupakan jenis diabetes yang muncul selama kehamilan dan berpotensi menyebabkan komplikasi serius bagi ibu dan janin. Mengidentifikasi pasien dengan risiko diabetes gestasional sejak dini sangat penting untuk intervensi yang efektif. Penelitian ini menggunakan pendekatan algoritma *Naïve Bayes* dalam klasifikasi data pasien untuk mendeteksi adanya diabetes gestasional. Algoritma *Naïve Bayes* dipilih karena kemampuannya dalam menangani data dengan fitur yang bervariasi serta proses klasifikasinya yang sederhana dan efisien. Data pasien yang digunakan mencakup beberapa variabel seperti usia, indeks massa tubuh (BMI), tekanan darah, dan riwayat kesehatan keluarga.Hasil uji menunjukkan bahwa model *Naïve Bayes* mampu mencapai tingkat akurasi yang memadai dalam mengklasifikasikan pasien dengan diabetes gestasional. Setelah dilakukan pengujian menggunakan validasi silang, algoritma ini menunjukkan akurasi tanpa seleksi fitur mencapai akurasi 77% dan dengan seleksi fitur mencapai 80%.Penambahan proses seleksi fitur juga dianalisis untuk melihat dampaknya terhadap performa model. Berdasarkan hasil penelitian ini, algoritma *Naïve Bayes* menunjukkan potensi yang baik sebagai alat bantu dalam mendeteksi diabetes gestasional, khususnya di lingkungan yang memerlukan pengolahan data cepat dan sederhana.

Kata Kunci: Diabetes Gestasional, Feature Selection, Klasifikasi, *Naïve Bayes*

Abstract - *Gestational diabetes is a type of diabetes that occurs during pregnancy and has the potential to cause serious complications for both mother and fetus. Early identification of patients at risk of gestational diabetes is crucial for effective intervention. This study utilizes the Naïve Bayes algorithm approach in classifying patient data to detect the presence of gestational diabetes. The Naïve Bayes algorithm was chosen due to its capability to handle data with diverse features and its simple, efficient classification process. The patient data used includes several variables such as age, body mass index (BMI), blood pressure, and family medical history.The test results indicate that the Naïve Bayes model achieves a satisfactory level of accuracy in classifying patients with gestational diabetes. After cross-validation testing, the algorithm achieved 77% accuracy without feature selection and 80% accuracy with feature selection. The addition of a feature selection process was also analyzed to assess its impact on model performance. Based on the results of this study, the Naïve Bayes algorithm shows strong potential as a tool for detecting gestational diabetes, especially in environments that require quick and simple data processing.*

Keywords*: Calssification, Feature Selection, Gestational Diabetes, Naïve Bayes.*

1. PENDAHULUAN

Menurut Federasi Diabetes Internasional (IDF), diabetes adalah kondisi kronis yang ditandai oleh kadar gula darah yang tinggi dalam tubuh. Jika kadar gula darah pada penderita diabetes tidak dikelola dengan baik, hal ini dapat memicu komplikasi serius yang mempengaruhi berbagai organ, seperti mata, ginjal, dan saraf, serta meningkatkan risiko penyakit jantung dan stroke. Faktor genetik, gaya hidup tidak sehat, serta perilaku tertentu merupakan faktor penyebab utama diabetes. Selain itu, faktor lingkungan sosial dan akses terhadap layanan kesehatan juga mempengaruhi terjadinya diabetes. Ada beberapa jenis diabetes, termasuk Diabetes tipe 1 (tergantung pada insulin), Diabetes tipe 2 (tidak tergantung pada insulin), serta tipe lain seperti diabetes gestasional (diabetes selama kehamilan).

Diabetes mellitus gestasional (DMG) adalah gangguan toleransi glukosa yang pertama kali terdeteksi pada wanita hamil. Kondisi ini biasanya dialami oleh wanita yang sebelumnya tidak didiagnosis dengan diabetes dan kemudian mengalami peningkatan kadar glukosa selama kehamilan. Penyebab utama dari diabetes ini adalah perubahan hormon. Selama kehamilan, plasenta memproduksi hormon seperti hormon lactogen plasenta (HPL), hormon estrogen, dan hormon yang dapat mengurangi sensitivitas tubuh terhadap insulin. Hormon-hormon ini semakin meningkat seiring dengan perkembangan kehamilan, yang berdampak pada kerja insulin. Biasanya, kondisi ini

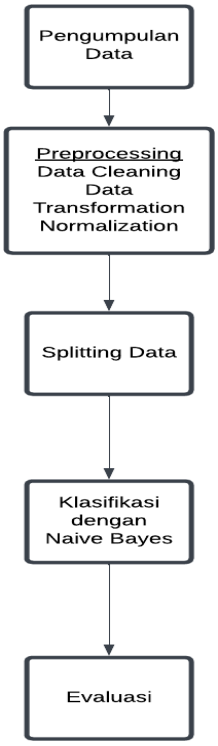
JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096

membaik setelah melahirkan ketika kadar gula darah kembali normal.Teknik data mining sering diterapkan untuk membantu proses diagnosis. Data mining adalah proses pengolahan data dalam jumlah besar untuk menemukan pola-pola tertentu yang bermanfaat, terutama dalam konteks medis. Teknik ini memungkinkan kita untuk menganalisis data pasien guna mengidentifikasi pola atau prediksi tertentu terkait kesehatan mereka.Beberapa indikator kesehatan dapat menunjukkan adanya penyakit diabetes dalam tubuh. Namun, keterbatasan tenaga medis dan teknologi membuat pengolahan data secara manual menjadi kurang efisien, terutama karena volume data yang sangat besar. Oleh karena itu, penerapan teknologi di bidang matematika sangat mendukung analisis di bidang medis. Salah satu teknik yang banyak diterapkan adalah *classification*, atau pengklasifikasian data. Dalam penelitian ini, metode yang digunakan untuk klasifikasi adalah *Naïve Bayes*.

Naïve Bayes adalah algoritma klasifikasi yang berdasarkan teorema probabilitas Bayes untuk melakukan prediksi. Algoritma ini menggambarkan hubungan antara beberapa kejadian menggunakan probabilitas kondisional, sehingga dapat melakukan klasifikasi berdasarkan fitur atau atribut pada suatu data. Algoritma ini memanfaatkan metode probabilitas untuk membuat prediksi dari data historis yang telah ada.

2. METODE PENELITIAN

Metode penelitian adalah kerangka kerja yang menjelaskan rangkaian aktivitas dalam studi ini, dimulai dari fase awal sampai fase akhir untuk mencapai sasaran yang telah ditetapkan. Alur proses penelitian ditampilkan seperti pada Gambar 1.



Gambar 1. Flowchart Metode Penelitian

2.1 Pengumpulan Data

Langkah awal dalam penelitian ini dimulai dengan kegiatan pengumpulan data. Dataset dapat diperoleh dari beragam sumber, dengan salah satu cara termudah adalah mengakses data melalui platform online yang menyediakan dataset untuk umum. Dalam studi ini, kumpulan data diabetes

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096

diunduh melalui platform Kaggle, sebuah situs yang menyediakan akses ke berbagai dataset publik. Dataset yang digunakan memiliki total 3526 entri dengan 17 parameter input yang terdiri dari:

- Class Label
- Case Number
- Age
- No Of Pregenacy
- Gestation
- BMI
- HDL
- Family History
- Unexplained
- PCOS
- Sys BP
- Dia BP
- OGTT
- Sedentary Lifestyle
- Prediabetes
- Hemoglobin

Serta 1 parameter output berupa hasil diagnosis (Outcome) yang menunjukkan status diabetes pasien.

2.2. Preprocessing

Preprocessing atau pra-pemrosesan adalah langkah awal yang dilakukan untuk mengolah data mentah sebelum melanjutkan ke tahap analisis lebih lanjut. Proses ini bertujuan untuk meningkatkan kualitas data sehingga analisis berikutnya menjadi lebih akurat. Meskipun dataset yang digunakan pada penelitian ini sudah berbentuk numerik, langkah-langkah preprocessing tetap perlu dijalankan untuk memastikan data sepenuhnya berupa angka, bebas dari gangguan, tidak mengandung nilai ekstrem, dan sudah distandardisasi. Oleh karena itu, tahap ini sangat penting untuk menjamin data siap diolah secara optimal. Beberapa langkah yang dilakukan dalam proses preprocessing pada penelitian ini meliputi:

a. Data cleaning

Pembersihan data atau data cleaning berfokus pada membersihkan dataset dari elemen-elemen yang tidak diperlukan, seperti data yang hilang (missing values), nilai-nilai pencilan (outliers), data yang duplikat, atau inkonsistensi lainnya. Tujuannya adalah agar data menjadi lebih siap untuk diolah dan dianalisis secara efektif.

b. Data Transformation

Transformasi data atau data transformation bertujuan untuk mengubah data mentah ke dalam format yang lebih sesuai untuk keperluan analisis. Ini bisa mencakup pengubahan skala, format, atau struktur data agar sesuai dengan kebutuhan penelitian.

c. Normalization

Normalisasi atau normalization adalah proses mengatur kembali nilai-nilai dalam dataset agar berada pada skala yang sama. Hal ini penting untuk memastikan bahwa perbandingan antar

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096

data menjadi lebih konsisten dan adil, terutama saat menggunakan metode analisis yang sensitif terhadap perbedaan skala data.

2.3 Splitting Data

Setelah melewati tahap Preprocessing, selanjutnya adalah tahap splitting data. Splitting data merupakan proses untuk memastikan bahwa model yang dibangun dapat digeneralisasi dengan baik pada data yang belum dilihat sebelumnya.

2.4 Klasifikasi *Naïve Bayes*

Setelah prose splitting data, langkah berikutnya dalam penelitian ini adalah membangun model klasifikasi. Dalam hal ini, model yang digunakan adalah *Naïve Bayes*. Algoritma *Naïve Bayes* merupakan teknik klasifikasi yang didasarkan pada Teorema Bayes, yang memanfaatkan prinsip-prinsip probabilitas untuk melakukan prediksi. Algoritma ini dikembangkan dari teori yang diusulkan oleh Thomas Bayes, seorang ilmuwan asal Inggris, yang bertujuan untuk memperkirakan peluang di masa depan dengan menggunakan informasi atau data dari kejadian sebelumnya.

Metode *Naïve Bayes* terkenal karena pendekatan probabilitas yang sederhana namun efektif, serta kemampuannya dalam memberikan hasil yang cepat dan akurat. Keunggulannya terletak pada kemudahan implementasi, efisiensi dalam menangani data berukuran besar, dan interpretasi hasil yang mudah dipahami. Itulah mengapa algoritma ini sering digunakan untuk berbagai aplikasi klasifikasi, terutama pada analisis prediktif yang memerlukan keputusan cepat dan akurat.

2.5 Proses Evaluasi

Proses evaluasi jurnal-jurnal dilakukan dengan menilai kualitas metodologi penelitian, kejelasan penyajian data, dan kesesuaian kesimpulan dengan hasil penelitian. Setiap artikel yang dipilih akan ditinjau secara kritis oleh tim peneliti untuk memastikan validitas dan reliabilitas informasi yang diperoleh. Selain itu, artikel-artikel tersebut akan dibandingkan untuk mengidentifikasi kesamaan, perbedaan, dan kontribusi unik dari masing-masing penelitian terhadap topik yang dikaji.

Tabel 1. Karakteristik Artikel Yang Dianalisa

No	Nama Peneliti dan Tahun	Metode yang dibahas	Tujuan penelitiannya	Hasil yang didapat
1	Yulia Darmi dan Alda Setiarina (2020)	*Naïve Bayes* dengan Feature Selection Information Gain	Memprediksi penyakit diabetes menggunakan kombinasi *Naïve Bayes* dengan feature selection untuk meningkatkan akurasi prediksi	Mencapai akurasi 77.34% setelah menggunakan feature selection, meningkat dari 75.78% tanpa feature selection
2	Rifki Indra Perwira dan Anastasia Trifiana (2019)	*Naïve Bayes*	Mengembangkan sistem prediksi diabetes berbasis web menggunakan *Naïve Bayes*	Tingkat akurasi mencapai 81.25% dari dataset yang digunakan
3	Muhammad Syafii dan Dedy Rahman Wijaya (2018)	*Naïve Bayes* yang dioptimasi dengan	Meningkatkan akurasi klasifikasi diabetes dengan mengoptimalkan	Akurasi meningkat menjadi 84.37% setelah optimasi, dibandingkan

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096

		Algoritma Genetika	parameter *Naïve Bayes*	dengan 77.21% menggunakan *Naïve Bayes* standar
4	Sulastri dan Muhammad Rizki (2021)	*Naïve Bayes*	Membangun sistem diagnosis awal diabetes mellitus berbasis aplikasi	Mencapai tingkat akurasi 76.92% dalam mendiagnosis diabetes mellitus
5	Ade Mubarok dan Setia Nugroho (2019)	Perbandingan antara C4.5 dan *Naïve Bayes*	Membandingkan efektivitas algoritma C4.5 dan *Naïve Bayes* dalam prediksi diabetes	*Naïve Bayes* menunjukkan performa lebih baik dengan akurasi 75.13% dibanding C4.5 yang mencapai 73.82%

3. ANALISA DAN PEMBAHASAN

3.1 Proses Design Dengan Rapid Minner

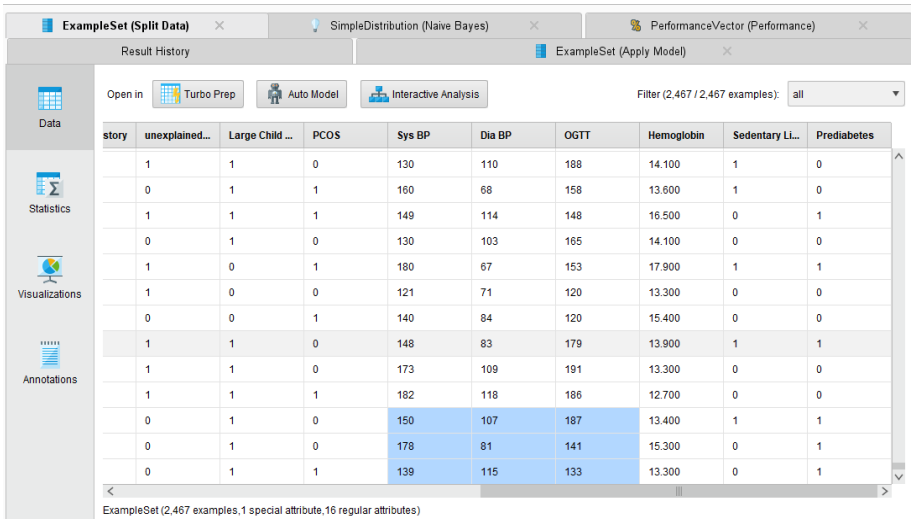
a. Pemrosesan Data

Rapid Miner menyediakan lingkungan design yang user-friendly untuk membangun alur kerja data mining yang kompleks. Proses design dimulai dengan membaca data, baik dari file lokal, database, atau sumber online. Data kemudian dipreproses untuk memastikan kualitas dan konsistensinya. Hal ini dapat melibatkan langkah-langkah seperti pembersihan data, transformasi data, dan penanganan missing values.

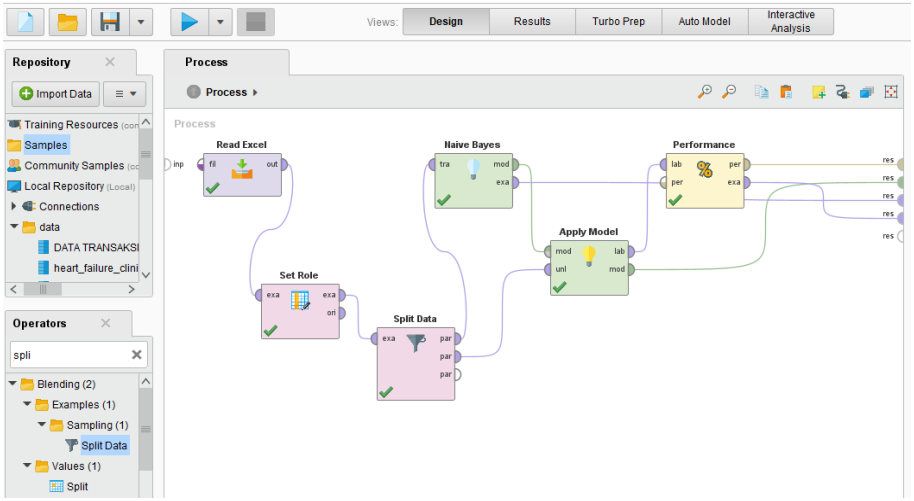
Setelah data siap, operator data mining diterapkan untuk melakukan analisis yang diinginkan. Rapid Miner menyediakan berbagai operator untuk berbagai tugas, seperti klasifikasi, clustering, regresi, dan association rule mining. Operator-operator ini dihubungkan dengan koneksi untuk membentuk alur kerja yang koheren. Alur kerja dapat dimodifikasi dan dioptimalkan dengan menggunakan fitur visualisasi yang intuitif.Hasil analisis dapat divisualisasikan dalam berbagai format, seperti tabel, grafik, dan peta. RapidMiner juga memungkinkan penyimpanan model dan deployment ke aplikasi real world.

Secara keseluruhan, proses design dengan RapidMiner memungkinkan pengguna untuk membangun alur kerja data mining yang efektif dan efisien, mulai dari pembacaan data hingga deployment model. Kemudahan penggunaan dan fleksibilitasnya menjadikannya alat yang ideal bagi para data scientist dan pemula dalam bidang data mining.

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096



Gambar 2. Dataset



Gambar 3. Desain Proses Rapid Miner

Gambar tersebut menunjukkan proses analisis data di RapidMiner yang bertujuan untuk membuat model prediksi menggunakan dataset yang berkaitan dengan diabetes gestasional.

Pertama, data diimpor dari file Excel menggunakan operator Read Excel, yang berfungsi untuk membaca data mentah ke dalam RapidMiner. Data yang diimpor kemudian diberikan peran khusus pada atribut-atributnya menggunakan operator Set Role. Dalam langkah ini, atribut yang menjadi target prediksi (label) ditetapkan, sementara atribut lainnya menjadi variabel prediktor.

Setelah itu, dataset dibagi menjadi dua bagian dengan operator Split Data untuk keperluan pelatihan (training) dan pengujian (testing) model. Pembagian data ini penting agar model dapat dilatih dengan sebagian data dan dievaluasi dengan data yang belum pernah dilihat sebelumnya untuk menghindari overfitting.

Model *Naïve Bayes* kemudian diterapkan pada data pelatihan. Algoritma ini dipilih karena kesederhanaannya dan kemampuannya dalam menangani data kategorikal, yang sering muncul pada dataset kesehatan. Setelah model dilatih, operator Apply Model digunakan untuk memprediksi data pada set pengujian.

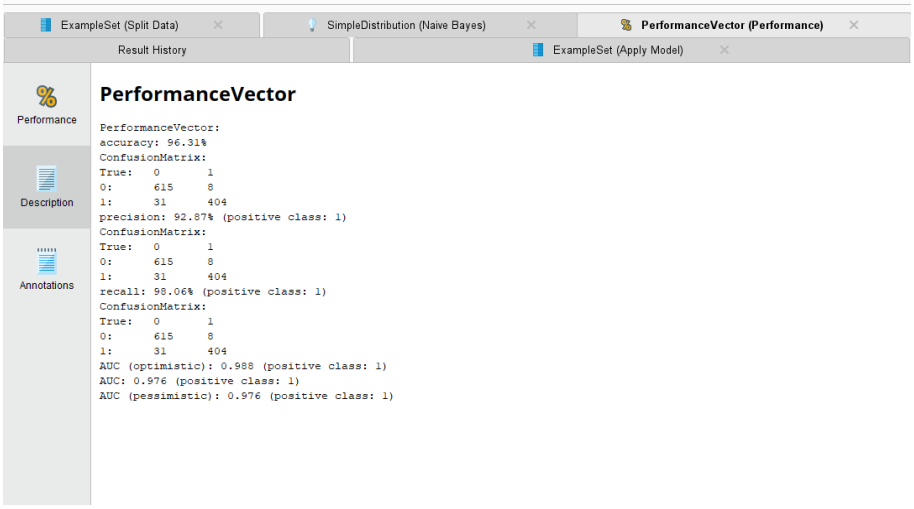
Langkah terakhir melibatkan operator Performance, yang digunakan untuk mengukur kinerja model yang telah dilatih. Evaluasi ini dilakukan dengan menghitung metrik seperti akurasi, presisi,

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096

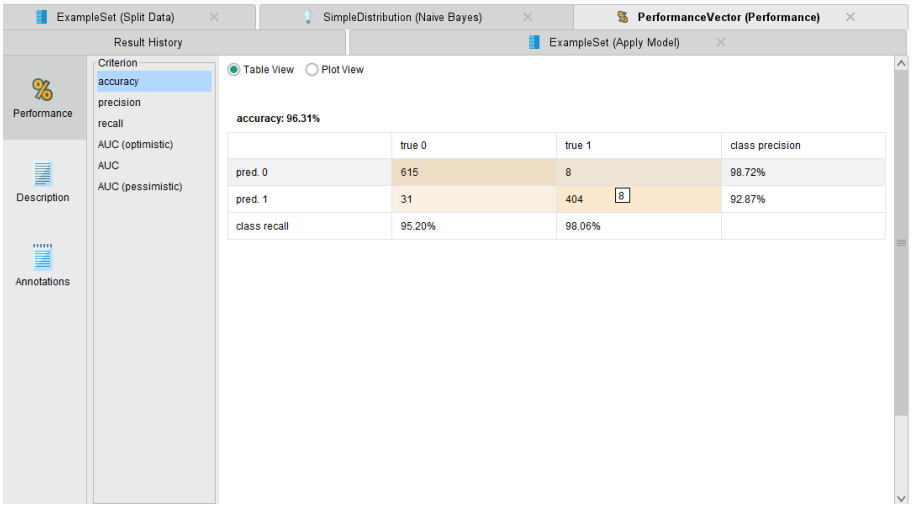
recall, dan F1-score untuk memastikan model bekerja dengan baik dalam memprediksi kasus diabetes gestasional pada data uji.

Secara keseluruhan, proses ini dirancang untuk membangun model prediksi yang dapat membantu tenaga medis dalam mengidentifikasi risiko diabetes gestasional berdasarkan data kesehatan pasien.

b. Performance



Gambar 4. *Performance Vector*



Gambar 5. *Accuracy*

Gambar menampilkan hasil evaluasi model di RapidMiner menggunakan dataset terkait diabetes gestasional. Berdasarkan hasil yang ditunjukkan pada confusion matrix dan metrik evaluasi, berikut adalah penjelasan mengenai akurasi serta metrik lain yang ditampilkan:

Penjelasan Hasil Evaluasi:

- Akurasi (*Accuracy*): 96.31%
- Akurasi adalah rasio jumlah prediksi yang benar terhadap total jumlah prediksi yang dibuat. Dalam hal ini, model berhasil mengklasifikasikan 96.31% dari semua sampel dengan benar. Ini berarti model *Naïve Bayes* yang digunakan memiliki performa yang sangat baik dalam memprediksi apakah pasien memiliki diabetes gestasional atau tidak.

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096

Penjelasan *Confusion Matrix*:

a. True 0 (Negatif Sebenarnya):
   - 615 sampel diklasifikasikan dengan benar sebagai kelas "0" (tidak memiliki diabetes gestasional).
   - 31 sampel diklasifikasikan salah sebagai "1" padahal seharusnya "0" (false positives).

b. True 1 (Positif Sebenarnya):
   - 404 sampel diklasifikasikan dengan benar sebagai kelas "1" (memiliki diabetes gestasional).
   - 8 sampel diklasifikasikan salah sebagai "0" padahal seharusnya "1" (false negatives).

Metrik Lainnya:

a. *Class Precision* (Presisi)
   - Untuk kelas "0": 98.72% (tingkat akurasi prediksi negatif).
   - Untuk kelas "1": 92.87% (tingkat akurasi prediksi positif).
   - Presisi mengukur seberapa akurat model saat memprediksi kelas tertentu, yaitu dari semua yang diprediksi sebagai positif, berapa banyak yang benar-benar positif.

b. *Class Recall* (Recall)
   - Untuk kelas "0": 95.20% (berapa banyak yang benar-benar negatif yang berhasil diprediksi).
   - Untuk kelas "1": 98.06% (berapa banyak yang benar-benar positif yang berhasil diprediksi).
   - Recall menunjukkan seberapa baik model dapat menemukan semua kasus positif atau negatif yang sebenarnya.

Model *Naïve Bayes* yang digunakan pada dataset diabetes gestasional menunjukkan kinerja yang sangat baik dengan akurasi tinggi (96.31%). Ini berarti model ini cukup andal untuk digunakan sebagai alat bantu diagnosis. Namun, perlu diperhatikan bahwa meskipun akurasi tinggi, keseimbangan antara presisi dan recall harus dipertimbangkan untuk memastikan model tidak terlalu bias ke salah satu kelas, terutama pada kasus medis di mana kesalahan klasifikasi bisa berdampak serius.

4. KESIMPULAN

Berdasarkan hasil evaluasi yang ditampilkan pada proses RapidMiner, model *Naïve Bayes* menunjukkan performa yang sangat baik dalam memprediksi risiko diabetes gestasional dengan akurasi mencapai 96.31%. Ini menunjukkan bahwa model dapat diandalkan dalam memberikan prediksi yang akurat berdasarkan data kesehatan pasien. Keberhasilan model ini juga ditunjukkan oleh presisi yang tinggi, yaitu 98.72% untuk kategori negatif (pasien yang tidak memiliki diabetes gestasional) dan 92.87% untuk kategori positif (pasien dengan diabetes gestasional). Hal ini berarti bahwa model mampu memprediksi dengan sangat akurat ketika menyatakan bahwa seorang pasien tidak memiliki kondisi tersebut, sekaligus memiliki keakuratan yang baik dalam mendeteksi pasien yang benar-benar berisiko.

Selain itu, recall untuk kategori positif mencapai 98.06%, menunjukkan kemampuan model yang kuat dalam mendeteksi kasus-kasus diabetes gestasional secara tepat. Meskipun model ini sangat efektif, terdapat beberapa kesalahan klasifikasi, yaitu sebanyak 31 kasus di mana pasien yang sebenarnya tidak memiliki diabetes gestasional diklasifikasikan sebagai positif (false positives), dan 8 kasus di mana pasien yang benar-benar memiliki kondisi tersebut diklasifikasikan sebagai negatif (false negatives). Meski jumlah kesalahan ini relatif kecil, terutama pada kasus false negatives, perlu

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 6, November Tahun 2024
ISSN 3025-0919 (media online)
Hal 1088-1096

diperhatikan mengingat dampak serius yang dapat terjadi jika seorang pasien yang memerlukan perawatan tidak terdeteksi.

Kesimpulan dari analisis ini adalah bahwa model berbasis *Naïve Bayes* dapat digunakan sebagai alat bantu diagnosis yang efektif untuk deteksi awal diabetes gestasional, memberikan peluang bagi intervensi dini yang dapat meningkatkan hasil kesehatan ibu hamil. Namun, untuk implementasi yang lebih luas di lingkungan klinis, disarankan agar penelitian lebih lanjut dilakukan dengan mengeksplorasi teknik machine learning lainnya dan pengujian pada dataset yang lebih beragam guna memastikan bahwa model ini dapat diaplikasikan secara umum pada berbagai populasi dan kondisi klinis.

REFERENCES

Darmi, Y., & Setiarina, A. (2020). *Prediksi penyakit diabetes menggunakan algoritma Naïve Bayes dengan feature selection Information Gain*. Jurnal Informatika dan Rekayasa Perangkat Lunak, 2(1), 78-86.

Perwira, R. I., & Trifiana, A. (2019). *Implementasi algoritma Naïve Bayes untuk prediksi penyakit diabetes.* Jurnal Teknologi Informasi dan Komunikasi, 10(2), 95-103.

Syafii, M., & Wijaya, D. R. (2018). *Klasifikasi penyakit diabetes mellitus menggunakan Naïve Bayes dengan optimasi parameter menggunakan algoritma genetika.* Jurnal Teknologi Informasi dan Ilmu Komputer, 5(4), 427-434.

Sulastri, & Rizki, M. (2021). *Penerapan metode Naïve Bayes untuk diagnosa penyakit diabetes mellitus*. Jurnal Teknik Informatika dan Sistem Informasi, 7(3), 512-521.

Mubarok, A., & Nugroho, S. (2019). *Komparasi kinerja algoritma C4.5 dan Naïve Bayes untuk prediksi penyakit diabetes.* Jurnal Ilmiah Teknologi Informasi Terapan, 5(2), 23-31.

Delvika, B., Nurhidayarnis, S., Rinada, P. D., Abror, N., & Hidayat, A. (2022). *Comparison of classification between Naïve Bayes and K-Nearest Neighbor on diabetes risk in pregnant women perbandingan klasifikasi antara Naïve Bayes dan K-Nearest Neighbor terhadap resiko diabetes pada ibu hamil.* MALCOM: Indonesian Journal of Machine Learning and Computer Science, 2(1), 68-75. Retrieved from https://journal.irpi.or.id/index.php/malcom/issue/view/17

Yolanda, V., Cholissodin, I., & Adikara, P. P. (2021). *Klasifikasi diagnosis penyakit diabetes gestasional pada ibu hamil menggunakan algoritme Neighbor Weighted K-Nearest Neighbor (NWKNN).* Retrieved from http://jptiik.ub.ac.id

Kaggle. (n.d.). Retrieved from https://www.kaggle.com

Universitas Pamulang Repository. (n.d.). Retrieved from http://repository.unpam.ac.id/id/eprint/12843

WDH Open Journal. (n.d.). Retrieved from https://openjournal.wdh.ac.id/index.php/JAM/article/view/134

Universitas Pamulang Repository. (n.d.). Retrieved from https://repository.unpam.ac.id/11847/