

Klasifikasi Kadar Glukosa Darah Keluaran Alat Non-invasif Menggunakan Regresi Logistik Ordinal dengan Peringkasan Luas *

Yuniar Istiqomah¹, Erfiani^{1‡}, and Utami Dyah Syafitri¹

¹Department of Statistics, IPB University, Indonesia

[‡]corresponding author: erfiani@apps.ipb.ac.id

Copyright © 2023 Yuniar Istiqomah, Erfiani, and Utami Dyah Syafitri. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Diabetes Mellitus (DM) is the silent killer because its symptoms tend to go unnoticed. The IPB Non-Invasive Biomarking Team developed a non-invasive monitoring device to check blood. The tool uses the spectroscopy principle and produces an output in the form of a residual value of light intensity. A method is needed to predict the category of blood glucose levels based on the measurement results of non-invasive tools. Classification modeling is one of the methods that can be used to analyze the relationship between the blood glucose level class of invasive measurement results and the residual value of the intensity of non-invasive measurement results. One of the commonly used classification methods is ordinal logistic regression. Light spectrum-based data used as predictor X changes often provide changes that correlate with each other. The principal component analysis reduces its dimensions to become a new set of changes that do not correlate. Graph area summation in the period is the best summarization method because it can take advantage of the general data information. This study uses the ordinal logistic regression method as a modeling method by applying principal component analysis and graph area summation applied to 2017 data and 2019 data. Classification modeling in the 2017 data had a balanced accuracy value of 64,64%. Classification modeling in the 2019 data produced a balanced accuracy value of 57,57%. The design used in the 2017 tool and the 2019 tool is different, causing the residual intensity graph of the non-invasive measurement results to be read differently. The 2017 data model is better applied to homogeneous data and the 2019 data model is better applied to heterogeneous data.

Keywords: blood glucose levels, graph area summation, non-invasive tool, ordinal logistic regression, principal component analysis.

* Received: Oct 2022; Reviewed: Dec 2022; Published: Jan 2023

1. Pendahuluan

Diabetes Melitus (DM) merupakan kelompok penyakit metabolik dengan karakteristik hiperglikemia yang terjadi karena kelainan sekresi insulin, kerja insulin atau keduanya (Perkeni 2015). Gejalanya cenderung tidak disadari dan menimbulkan berbagai keluhan lainnya sehingga disebut *silent killer*. Pemeriksaan kadar glukosa darah secara berkala penting dilakukan oleh penderita maupun non-penderita DM. Metode pemeriksaan kadar glukosa darah yang paling umum dilakukan adalah metode invasif. Metode ini dilakukan menggunakan alat glukometer atau melalui tes laboratorium. Metode invasif menimbulkan ketidaknyamanan karena melukai bagian tubuh sehingga dikembangkan alat pengukur kadar glukosa darah non-invasif yang tidak melukai tubuh.

Tim Non-Invasif IPB mengembangkan alat pengukur kadar glukosa darah non-invasif tanpa melukai tubuh dengan prinsip spektroskopi inframerah. Keluaran alat tersebut berupa nilai residu intensitas cahaya terhadap *time domain* sehingga diperlukan suatu metode untuk dapat memprediksi kategori kadar glukosa darah berdasarkan hasil pengukuran non-invasif tersebut. Pemodelan klasifikasi merupakan salah satu metode yang dapat digunakan untuk menganalisis hubungan antara kelas kadar glukosa darah hasil pengukuran invasif dengan nilai residu intensitas hasil pengukuran alat non-invasif. Klasifikasi bertujuan memprediksi kategori atau kelas dari suatu objek.

Kelas kadar glukosa darah invasif menurut ADA (2010) dibagi menjadi 3, yaitu tingkat kadar glukosa darah rendah (< 70 mg/dL), normal (70 mg/dl – 99 mg/dL) dan tinggi (≥ 99 mg/dL). Menurut ADA (2018), kelas kadar glukosa darah invasif juga dapat dibagi menjadi 3 kategori, yaitu tingkat kadar glukosa darah normal (< 100 mg/dL), prediabetes (100 mg/dL – 126 mg/dL) dan diabetes (≥ 126 mg/dL). Kelas kadar glukosa darah tersebut berupa kategori yang bersifat tingkatan.

Salah satu metode klasifikasi yang umum digunakan untuk menghubungkan antara peubah bebas berupa numerik atau kategorik dan peubah respon berupa dua atau lebih kategori bersifat ordinal adalah regresi logistik ordinal (Hosmer dan Lemeshow 2000). Metode ini menghasilkan model akhir berupa fungsi logit peluang kumulatif (Agresti 2002). Data berbasis spektrum cahaya yang digunakan sebagai peubah prediktor X seringkali memberikan peubah yang saling berkorelasi antara satu dengan lainnya (Aurelia 2020). Hal tersebut akan diatasi menggunakan analisis komponen utama (*principal component analysis*) untuk mereduksi dimensi peubah hingga menjadi sekumpulan peubah baru yang tidak berkorelasi.

Penelitian sebelumnya mengenai pemodelan klasifikasi kelas kadar glukosa darah non-invasif dilakukan oleh Sari (2019). Sari (2019) menggunakan metode regresi logistik ordinal pada data tidak seimbang dengan pendekatan peringkasan rataan dan simpangan baku. Penelitian lainnya juga dilakukan oleh Annisa (2019). Annisa (2019) menggunakan metode *Support Vector Machine* (SVM) dengan pendekatan peringkasan tiga nilai, peringkasan rataan, dan peringkasan simpangan baku. Kedua penelitian tersebut menunjukkan bahwa peringkasan simpangan baku menghasilkan ketepatan klasifikasi paling baik dibandingkan dengan peringkasan lainnya. Namun demikian, peringkasan simpangan baku tidak memanfaatkan seluruh informasi data.

Peringkasan yang digunakan untuk mengatasi kekurangan dari peringkasan simpangan baku adalah peringkasan luas. Peringkasan luas ini memanfaatkan nilai residu intensitas dan *time domain*. Konsep peringkasan luas dilakukan dengan perhitungan luas grafik residu intensitas dengan pendekatan luas trapesium. Penelitian sebelumnya yang menggunakan peringkasan luas, diantaranya Aurelia (2020), Satrianto (2020), dan Wibawati (2021). Aurelia 2020 menyatakan bahwa peringkasan luas dapat memberikan pendugaan kadar glukosa darah dengan performa terbaik. Satrianto (2020) menyatakan bahwa peringkasan luas memberikan nilai rata-rata MSE lebih kecil. Wibawati (2021) menyebutkan bahwa peringkasan luas memberikan ketepatan klasifikasi paling baik. Berdasarkan hal tersebut, penelitian ini akan menyusun model klasifikasi kadar glukosa darah keluaran alat non-invasif yang dikembangkan tim non-invasif biomarking IPB menggunakan regresi logistik ordinal dengan menerapkan analisis komponen utama (*principal component analysis*) dan peringkasan luas serta membandingkan performa model pada data 2017 dan 2019.

2. Metodologi

2.1 Data

Data yang digunakan dalam penelitian ini adalah data hasil penelitian “Model Kalibrasi Melalui Model Statistik Transformasi Wavelet Kontinu, Regresi Kuadrat Terkecil Parsial Robust, Regresi Proses Gaussian untuk Pengembangan Sistem Alat Pemantau Pasien Penderita Kadar Gula Darah Tinggi secara Non-Invasif” oleh Tim Non-Invasif Biomarking IPB pada tahun 2017 dan 2019. Pengambilan data periode April 2016 – Januari 2017 melibatkan 118 responden yang merupakan mahasiswa IPB. Selanjutnya, kelompok data ini disebut sebagai data 2017. Pengambilan data periode Juli 2019 melibatkan 74 responden yang merupakan warga Kelurahan Kebon Pedes Kecamatan Tanah Sereal Kota Bogor. Kelompok data ini disebut data 2019. Nilai residu intensitas hasil pengukuran alat non-invasif akan digunakan sebagai peubah prediktor dan kelas kadar glukosa darah hasil pengukuran invasif dengan metode *gold standard* akan digunakan sebagai peubah respon.

2.2 Prosedur Analisis Data

Prosedur analisis data yang digunakan masing-masing data 2017 dan data 2019 pada penelitian ini adalah sebagai berikut:

1. Mengkategorikan data kadar glukosa darah hasil pengukuran invasif menjadi 3 kategori.
 - a. Kategori yang digunakan pada data 2017 menurut ADA (2010), yaitu menggunakan tingkat kadar glukosa darah rendah (< 70 mg/dL), normal (70 mg/dl – 99 mg/dL) dan tinggi (≥ 99 mg/dL).
 - b. Kategori yang digunakan pada data 2019 menurut ADA (2018), yaitu menggunakan tingkat kadar glukosa darah normal (< 100 mg/dL), prediabetes (100 mg/dL – 126 mg/dL) dan diabetes (≥ 126 mg/dL).
2. Melakukan prapemrosesan data hasil pengukuran kadar glukosa darah non-invasif.
 - a. Peringkasan luas menggunakan pendekatan luas trapesium. Persamaan pendekatan peringkasan luas adalah sebagai berikut:

$$XL_k = \frac{1}{2} \sum_{i=1}^{n-1} (t_{i+1} - t_i)(y_i + y_{i+1}) \quad (1)$$

$k = 1, 2, \dots, 25$ (berdasarkan periode)

dengan XL_k menyatakan luas grafik residu intensitas periode ke- k , t_i merupakan *time domain* ke- i , y_i adalah nilai residu intensitas ke- i , dan n adalah banyak titik batas *time domain* pada periode tersebut.

- b. Melakukan analisis komponen utama dan menentukan jumlah komponen utama yang akan digunakan sebagai peubah prediktor berdasarkan besarnya nilai keragaman kumulatif. Penentuan jumlah Komponen Utama (KU) yang akan digunakan untuk analisis selanjutnya dilakukan dengan cara melihat keragaman kumulatif yang dapat dijelaskan lebih dari 80% (Johnson dan Winchern 2007).
3. Melakukan analisis statistika deskripsi data kadar glukosa darah hasil pengukuran invasif dan non-invasif.
4. Mendeteksi adanya multikolinearitas menggunakan matriks korelasi dan mendeteksi adanya pencilan. Pencilan adalah suatu amatan yang titik pengamatannya menyimpang dari pola data. Pencilan dapat disebabkan oleh diantaranya kesalahan alat ukur dan kesalahan saat pengambilan data. Penelitian ini menggunakan Jarak Mahalanobis untuk mendeteksi adanya pencilan. Metode Jarak Mahalanobis menggunakan vektor rata-rata dan matriks varian-kovarian (Sari *et al.* 2021). Jarak Mahalanobis yang diperoleh kemudian disajikan dalam bentuk diagram kotak garis untuk mendeteksi adanya pencilan. Persamaan jarak malahanobis adalah sebagai berikut:

$$D_M(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)} \quad (2)$$

dengan x adalah matriks peubah prediktor, Σ adalah matiks kovarian, dan μ adalah vektor rata-rata.

5. Membagi data baru yaitu data dengan peubah prediktor berupa skor komponen utama yang terpilih menjadi gugus data latih utama dan data uji utama dengan proporsi 80% data latih utama dan 20% data uji utama secara acak. Selanjutnya, data baru juga dibagi meggunakan validasi silang *k-fold* dengan $k = 5$.
6. Melakukan SMOTE pada data latih utama untuk menangani kelas yang tidak seimbang.
7. Melakukan pemodelan menggunakan regresi logistik ordinal.
8. Mengevaluasi kinerja model dengan menghitung nilai *balanced accuracy* model yang diperoleh.
9. Mentransformasi model terpilih ke bentuk peubah asalnya yaitu peubah X.

3. Hasil dan Pembahasan

3.1 Prapemrosesan Data Hasil Pengukuran Non-Invasif

3.1.1 Peringkasan Luas

Peringkasan luas dilakukan agar jumlah titik residu intensitas yang dihasilkan pada setiap modulasi atau periode seragam dengan *time domain* yang konsisten. Peringkasan luas dilakukan dengan menjumlahkan luas grafik residu intensitas yang berada dalam satu periode melalui pendekatan luas trapesium. Perhitungannya

melibatkan komponen nilai residu intensitas serta *time domain* dari data agar peringkasan tidak membuang terlalu banyak data. Peringkasan dilakukan pada masing-masing data dalam satu periode yang diulang sebanyak 5 kali ulangan sehingga diperoleh sebuah angka untuk mewakili nilai dari satu periode berupa nilai luas grafik. Peringkasan ini akan menghasilkan 25 peubah prediktor.

3.1.2 Analisis Komponen Utama

Analisis komponen utama menghasilkan skor Komponen Utama (KU) yang akan digunakan untuk analisis selanjutnya. Penelitian ini akan menggunakan jumlah KU dengan keragaman kumulatif >90%. Data 2017 akan menggunakan jumlah KU sebanyak 5 karena keragaman kumulatif yang dapat dijelaskan sebesar 90,139% sedangkan data 2019 akan menggunakan jumlah KU sebanyak 4 karena keragaman yang dapat dijelaskan sebesar 91,278%. Selanjutnya, skor komponen utama yang terpilih akan digunakan sebagai peubah prediktor.

3.2 Eksplorasi Kadar Glukosa Darah Invasif

Pengukuran kadar glukosa darah invasif pada tahun 2017 melibatkan 118 responden. Tabel 1 menggambarkan kadar glukosa darah terkecil pada responden data 2017 sebesar 67 mg/dL dan kadar glukosa darah tertinggi sebesar 276 mg/dL dengan rata-rata kadar glukosa darah secara keseluruhan sebesar 82,64 mg/dL. Pengukuran kadar glukosa darah pada tahun 2019 melibatkan 74 responden. Kadar glukosa darah terkecil responden pada data 2019 sebesar 69 mg/dL dan kadar glukosa darah tertinggi sebesar 614 mg/dL dengan rata-rata kadar glukosa darah secara keseluruhan sebesar 140,7 mg/dL. Data 2019 memiliki nilai keragaman yang lebih besar dibandingkan dengan data 2017. Hal tersebut dikarenakan responden yang terpilih pada tahun 2019 memiliki rentang usia lebih lebar dan karakteristik kadar glukosa darah lebih beragam daripada responden tahun 2017.

Tabel 1 Deskripsi hasil pengukuran kadar glukosa darah invasif tahun 2017 dan 2019

Tahun	n	Minimum	Median	Maksimum	Rata-rata	Simpangan Baku
2017	118	67	80	276	82,64	20,4
2019	74	69	105,5	614	140,7	87,7

Kadar glukosa darah hasil pengukuran invasif pada data 2017 juga dibagi menjadi tiga kelas berdasarkan kategori menurut ADA (2010), yaitu tingkat kadar glukosa darah rendah (< 70 mg/dL), normal (70 mg/dL – 99 mg/dL), dan tinggi (\geq 99 mg/dL). Tabel 2 menunjukkan jumlah responden pada data 2017 yang memiliki kadar glukosa darah normal sebesar 102 responden. Jumlah tersebut jauh lebih banyak dibandingkan dengan jumlah responden dengan kadar glukosa darah rendah dan tinggi dengan masing – masing jumlah responden sebesar 8 responden. Oleh karena itu, data tersebut termasuk dalam kategori kelas tidak seimbang. Kelas kadar glukosa darah normal disebut sebagai kelas mayor sedangkan kelas kadar glukosa darah rendah dan tinggi disebut kelas minor.

Tabel 2 Persentase kelas kadar glukosa darah invasif tahun 2017 dan 2019

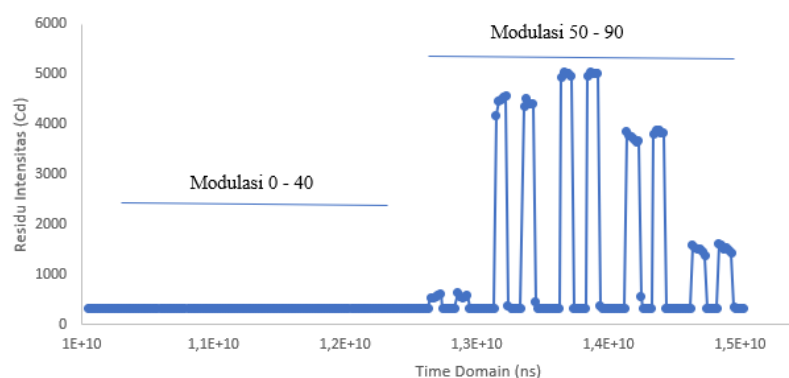
.Data 2017			Data 2019		
Kelas	Jumlah	Persentase (%)	Kelas	Jumlah	Persentase (%)
Rendah	8	6,78	Normal	34	45,9
Normal	102	86,44	Prediabetes	11	14,9
Tinggi	8	6,78	Diabetes	29	39,2
Total	118	100	Total	74	100

Kadar glukosa darah invasif tahun 2019 juga dibagi menjadi tiga kelas berdasarkan kategori menurut ADA (2018), yaitu tingkat kadar glukosa darah normal (< 100 mg/dL), prediabetes (100 mg/dL – 126 mg/dL), dan diabetes (≥ 126 mg/dL). Hal tersebut dikarenakan responden yang terpilih pada tahun 2019 memiliki rentang usia lebih lebar dan karakteristik kadar glukosa darah lebih beragam daripada responden tahun 2017. Tabel 2 menunjukkan 45,9% responden atau 34 responden memiliki kadar glukosa darah normal, 14,9% atau 11 responden dalam kategori prediabetes dan 39,2% atau 29 responden dalam kategori diabetes.

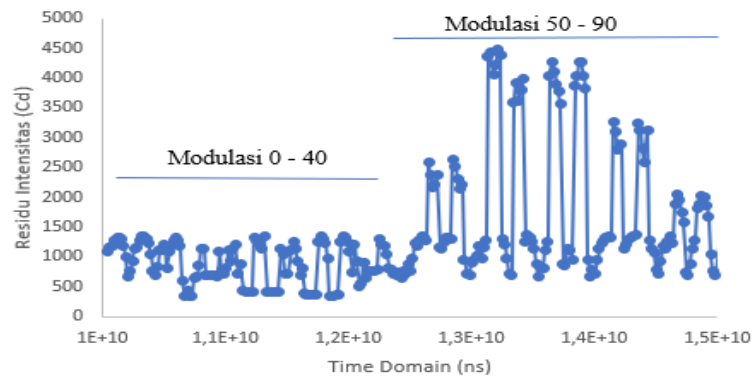
3.3 Eksplorasi Nilai Residu Intensitas Hasil Pengukuran Non-Invasif

Data Hasil pengukuran alat non-invasif berupa residu intensitas terhadap *time domain*. Data ini selanjutnya akan digunakan sebagai peubah prediktor. Data nilai residu intensitas yang akan digunakan adalah data pada periode 21 – 30. Hal tersebut mengacu pada penelitian Puteri (2018) yang menyebutkan bahwa lampu inframerah 1600 nm yang digunakan pada pengukuran periode 21 – 30 adalah lampu yang dikhususkan untuk terserap pada gula, sedangkan lampu dengan panjang gelombang 1550 nm dikhususkan untuk terserap pada air.

Pengambilan data 2017 menggunakan alat non-invasif dengan sensor yang didesain menangkap residu dari intensitas cahaya yang diserap oleh anggota tubuh. Gambar 1 menampilkan grafik pengukuran alat pada periode 21 – 30. Residu intensitas yang dihasilkan pada modulasi 0 sampai 40 tidak membentuk puncak dan cenderung konstan. Intensitas yang diberikan dari modulasi-modulasi tersebut hampir sepenuhnya diserap oleh senyawa tubuh. Residu intensitas yang dihasilkan pada modulasi 50 sampai 90 membentuk puncak yang artinya intensitas yang dihasilkan sampai ke sensor dan tidak diserap sepenuhnya oleh senyawa tubuh.



Gambar 1 Grafik keluaran alat non-invasif salah satu ulangan pada responden tahun 2017



Gambar 2 Grafik keluaran alat non-invasif salah satu ulangan pada responden tahun 2019

Pengambilan data 2019 menggunakan alat non-invasif dengan sensor yang didesain menangkap residu dari intensitas cahaya yang terpantul oleh anggota tubuh. Gambar 2 merupakan grafik keluaran alat non-invasif salah satu ulangan pada responden tahun 2019. Grafik tersebut terlihat lebih fluktuatif dibandingkan dengan data 2017. Hal ini diduga karena ada perbedaan desain, pengaturan pada alat serta sensor yang dinilai jauh lebih sensitif daripada alat 2017. Walaupun intensitas yang diperoleh lebih fluktuatif, modulasi 0 sampai 40 juga cenderung konstan dibandingkan dengan modulasi 50 sampai 90. Menurut Aurelia (2020), modulasi 50 sampai 90 pada data 2017 dan 2019 menunjukkan nilai residu yang signifikan dengan jumlah titik residu intensitas dan *time domain* yang terbaca berbeda. Penelitian ini akan menggunakan modulasi 50 – 90 (periode 26 – 30).

3.3.1 Deteksi Multikolinearitas

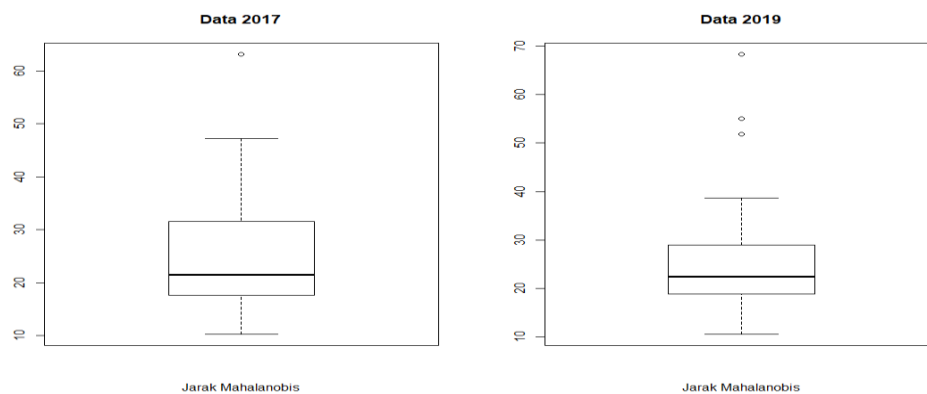
Multikolinearitas dapat diketahui dengan melihat matriks korelasi. Jika antar peubah X terdapat korelasi yang tinggi umumnya mendekati 1 atau -1, hal ini menjadi indikasi adanya multikolinearitas. Matriks korelasi antar peubah X pada data 2017 dan 2019 dapat dilihat pada Tabel 3.

Tabel 3 Matriks korelasi antar peubah X pada data 2017 dan data 2019

Data 2017					Data 2019				
	X1	X2	...	X25		X1	X2	...	X25
X1	1	0,72	...	0,78	X1	1	0,55	...	0,61
X2	0,72	1	...	0,78	X2	0,55	1	...	0,29
X3	0,71	0,82	...	0,82	X3	0,90	0,49	...	0,53
X4	0,72	0,79	...	0,87	X4	0,84	0,46	...	0,65
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
X25	0,78	0,78	...	1	X25	0,61	0,29	...	1

Tabel 3 menunjukkan adanya korelasi tinggi antar peubah karena nilai korelasi mendekati 1. Hal ini mengindikasikan terjadinya multikolinearitas antar peubah prediktor pada data 2017 dan data 2019. Pendekatan model klasifikasi yang akan digunakan adalah *principal component logistic regression* (PCLR).

3.3.2 Deteksi Pencilan



Gambar 3 Diagram kotak garis jarak mahalanobis pada data 2017 dan 2019

Pencilan adalah suatu data yang menyimpang dari data yang lain. Salah satu metode untuk mendeteksi pencilan adalah menggunakan Jarak Mahalanobis. Sebaran Jarak Mahalanobis pada data 2017 dan 2019 dapat dilihat pada Gambar 3. Gambar 3 menunjukkan bahwa terdapat pencilan baik pada data 2017 maupun 2019. Jarak mahalanobis yang termasuk dalam pencilan pada data 2017 adalah jarak mahalanobis pada responden ke-2, sedangkan yang termasuk pencilan pada data 2019 adalah responden ke-3, ke-11, dan ke-48. Pencilan ini tidak dihilangkan untuk menghindari kehilangan informasi data.

3.4 Penanganan Data Tidak Seimbang

Metode SMOTE digunakan untuk menangani data yang tidak seimbang. Metode ini diterapkan pada gugus data latih. Data 2017 dibagi menjadi 3 kelas dengan kelas data minor sebanyak 2 kelas. Setiap amatan kelas minoritas akan dibangkitkan 8 amatan kelas minoritas buatan. Kelas mayoritas pada data yaitu kelas kadar glukosa darah normal tidak dilakukan pengurangan jumlah data untuk menghindari berkurangnya informasi data pada kelas tersebut. Data 2019 juga dibagi menjadi 3 kelas dengan kelas data minor sebanyak 1 kelas. Setiap amatan kelas minoritas akan dibangkitkan 2 amatan kelas minoritas buatan.

3.5 *Principal Component Logistic Regression (PCLR)*

Model klasifikasi dengan metode regresi logistik ordinal dibentuk menggunakan data latih dengan skor komponen utama yang terpilih sebagai peubah prediktor. Performa hasil klasifikasi dari proses validasi silang *5-fold* pada data 2017 dan 2019 dapat dilihat pada Tabel 4. Penentuan model klasifikasi terbaik dilakukan dengan membandingkan performa model utama (model yang dibangun menggunakan data latih utama) dengan model *fold* yang dicobakan pada data uji utama. Tabel 4 menunjukkan model *fold* 1 pada data 2017 menghasilkan *balanced accuracy* tertinggi sehingga terpilih untuk dibandingkan dengan model utama. Model *fold* 1 pada data 2019 juga menghasilkan *balanced accuracy* tertinggi, sehingga terpilih untuk dibandingkan dengan model utama.

Tabel 4 Performa model klasifikasi dalam setiap fold pada data 2017 dan 2019

Data 2017		Data 2019	
Fold	Balance Accuracy (%)	Fold	Balanced Accuracy (%)
1	69,37	1	68,90
2	49,68	2	62,67
3	52,48	3	56,18
4	47,42	4	56,21
5	53,24	5	55,88

3.6 Perbandingan Model Klasifikasi Terbaik

Penentuan model klasifikasi terbaik dilakukan dengan membandingkan performa model utama (model yang dibangun menggunakan data latih utama) dengan model *fold* terbaik yang dicobakan pada data uji utama. Hasil evaluasi model *fold* terbaik terhadap data uji utama dapat dilihat pada Tabel 5. Model yang ditetapkan sebagai model terbaik pada data 2017 adalah model *fold* 1 karena menghasilkan nilai *balanced accuracy* yang lebih tinggi dibandingkan dengan model utama. Model yang ditetapkan sebagai model terbaik pada data 2019 adalah model utama karena menghasilkan nilai *balanced accuracy* yang lebih tinggi dibandingkan dengan model *fold* 1.

Tabel 5 Perbandingan performa model klasifikasi pada data 2017 dan 2019

Data 2017		Data 2019	
Model	Balanced Accuracy (%)	Model	Balanced Accuracy (%)
Model utama	36,98	Model utama	57,57
Model <i>fold</i> 1	64,64	Model <i>fold</i> 1	55,86

Model terbaik yang digunakan pada data 2017 adalah model *fold* 1 dengan nilai *balanced accuracy* sebesar 64,64%. Penduga parameter model tersebut dapat dilihat pada Tabel 6 dan menghasilkan persamaan model sebagai berikut:

$$\text{Logit } [P(Y \leq \text{Rendah}|x)] = -0,96 + 0,236PC1 + 0,165PC2 + 0,586PC3 - \\ 0,639PC4 - 1,556PC5$$

$$\text{Logit } [P(Y \leq \text{Normal}|x)] = 1,33 + 0,236PC1 + 0,165PC2 + 0,586PC3 - \\ 0,639PC4 - 1,556PC5$$

Tabel 6 Penduga parameter model regresi logistik ordinal data 2017

Peubah	Koefisien	Std. Error	t-value
Koef. Rendah	-0,960	0,359	-2,673
Koef. Normal	1,330	0,400	3,328
PC1	0,236	0,853	2,766
PC2	0,165	0,242	0,682
PC3	0,586	0,522	1,123
PC4	-0,639	0,614	-1,041
PC5	-1,556	0,648	-2,406

Model terbaik yang digunakan pada data 2019 adalah model utama dengan nilai *balanced accuracy* sebesar 57,57%. Penduga parameter model tersebut dapat dilihat pada tabel 7 dan menghasilkan persamaan model sebagai berikut:

$$\text{Logit} [P(Y \leq \text{Normal}|x)] = -0,808 - 0,138PC1 + 0,004PC2 + 1,672PC3 + 0,173PC4$$

$$\begin{aligned} \text{Logit} [P(Y \leq \text{Prediabetes}|x)] \\ = 0,713 - 0,138PC1 + 0,004PC2 + 1,672PC3 + 0,173PC4 \end{aligned}$$

Model terpilih kemudian ditransformasi ke bentuk peubah asalnya yaitu peubah X. Model regresi logistik ordinal setelah dilakukan transformasi adalah sebagai berikut: Model pada data 2017:

$$\begin{aligned} \text{Logit}[P(Y \leq \text{Rendah}|x)] = & -0,96 + 0,09 X_1 + 0,549 X_2 + 0,172 X_3 - \\ & 0,326 X_4 + 0,025 X_5 + 0,174 X_6 - 0,196 X_7 + \\ & 0,216 X_8 + 0,007 X_9 + 0,192 X_{10} - 0,271 X_{11} - \\ & 0,125 X_{12} - 0,094 X_{13} - 0,348 X_{14} + \\ & 0,231 X_{15} - 0,445 X_{16} - 0,105 X_{17} + \\ & 0,615 X_{18} + 0,251 X_{19} + 0,185 X_{20} - \\ & 0,274 X_{21} - 0,696 X_{22} - 0,16 X_{23} - \\ & 0,051 X_{24} + 0,25 X_{25} \end{aligned}$$

$$\begin{aligned} \text{Logit}[P(Y \leq \text{Normal}|x)] = & 1,33 + 0,09 X_1 + 0,549 X_2 + 0,172 X_3 - \\ & \mathbf{0,326 X_4 + 0,025 X_5 + 0,174 X_6 - 0,196 X_7 +} \\ & \mathbf{0,216 X_8 + 0,007 X_9 + 0,192 X_{10} - 0,271 X_{11} -} \\ & 0,125 X_{12} - 0,094 X_{13} - 0,348 X_{14} + \\ & 0,231 X_{15} - 0,445 X_{16} - 0,105 X_{17} + \\ & 0,615 X_{18} + 0,251 X_{19} + 0,185 X_{20} - \\ & 0,274 X_{21} - 0,696 X_{22} - 0,16 X_{23} - \\ & 0,051 X_{24} + 0,25 X_{25} \end{aligned}$$

Model pada data 2019:

$$\begin{aligned} \text{Logit}[P(Y \leq \text{Normal}|x)] = & -0,808 - 0,151 X_1 - 1,463 X_2 - 0,11 X_3 - \\ & 0,068 X_4 - 0,016 X_5 + 0,031 X_6 + 0,25 X_7 + \\ & 0,18 X_8 - 0,078 X_9 - 0,174 X_{10} + 0,176 X_{11} + \\ & 0,336 X_{12} - 0,077 X_{13} - 0,012 X_{14} - \\ & 0,056 X_{15} + 0,17 X_{16} + 0,097 X_{17} + \\ & 0,077 X_{18} + 0,066 X_{19} + 0,419 X_{20} + \\ & 0,117 X_{21} + 0,197 X_{22} + 0,113 X_{23} - \\ & 0,029 X_{24} + 0,25 X_{25} \end{aligned}$$

$$\begin{aligned} \text{Logit}[P(Y \leq \text{Prediabetes}|x)] = & 0,713 - 0,151 X_1 - 1,463 X_2 - 0,11 X_3 - \\ & 0,068 X_4 - 0,016 X_5 + 0,031 X_6 + 0,25 X_7 + \\ & 0,18 X_8 - 0,078 X_9 - 0,174 X_{10} + 0,176 X_{11} + \\ & 0,336 X_{12} - 0,077 X_{13} - 0,012 X_{14} - \\ & 0,056 X_{15} + 0,17 X_{16} + 0,097 X_{17} + \\ & 0,077 X_{18} + 0,066 X_{19} + 0,419 X_{20} + \\ & 0,117 X_{21} + 0,197 X_{22} + 0,113 X_{23} - \\ & 0,029 X_{24} + 0,25 X_{25} \end{aligned}$$

4 Simpulan dan Saran

4.1 Simpulan

Model terbaik yang diperoleh dari *Principal component logistic regression* (PCLR) pada data 2017 adalah model *fold* 1 dengan nilai *balanced accuracy* sebesar 64,64 % dan model terbaik pada data 2019 adalah model utama yang menghasilkan nilai *balanced accuracy* sebesar 57,57%. Desain yang digunakan pada alat 2017 dan alat 2019 berbeda. Hal ini menyebabkan grafik residu intensitas hasil pengukuran non-invasif yang terbaca juga berbeda. Model pada data 2017 lebih baik diterapkan pada data homogen dengan rentang kadar glukosa darah yang sempit dan model pada data 2019 lebih baik diterapkan pada data heterogen yaitu rentang kadar glukosa darah lebih lebar.

Model terbaik yang diperoleh dari *Principal component logistic regression* (PCLR) pada data 2017 adalah model *fold* 1 dengan nilai *balanced accuracy* sebesar 64,64% sedangkan model terbaik pada data 2019 adalah model utama yang menghasilkan nilai *balanced accuracy* sebesar 57,57%. Model terbaik pada data 2017 menghasilkan performa yang lebih tinggi dibandingkan dengan data 2019.

4.2 Saran

Perlu dilakukan kajian lanjutan mengenai metode peringkasan lainnya untuk mendapatkan peubah prediktor dari keluaran alat non-invasif. Pengkajian lanjutan mengenai batasan kadar glukosa darah yang digunakan juga perlu dilakukan serta dicobakan metode klasifikasi lain yang lebih baik dari metode regresi logistik ordinal dalam mengklasifikasikan data spektrum keluaran alat non-invasif dan dapat mengatasi data yang tidak seimbang.

Daftar Pustaka

- [ADA] American Diabetes Association. 2010. Position statement: Standards of Medical Care in Diabetes 2010. *Diab Care*. 2010: 33 (Suppl.1)
- [ADA] American Diabetes Association (USA). 2018. Factors Affecting Blood Glucose [Internet]. [diakses 2021 November 4]. Tersedia pada: <http://www.diabetes.org/living-with-diabetes/treatment-and-care/bloodglucose-control/factors-affecting-bloodglucose.html?referrer=https://www.google.co.id/>
- Agresti A. 2002. *Categorical Data Analilysis*. New York (US): John Wiley and Sons, Inc.
- Annisa S. 2019. Klasifikasi kadar glukosa darah keluaran alat non-invasif menggunakan metode support vector machine [tesis]. Bogor (ID): Institut Pertanian Bogor.

- Aurelia K. 2020. Pendugaan kadar glukosa darah non-invasif menggunakan regresi kuadrat terkecil parsial dengan beberapa pendekatan peringkasan [skripsi]. Bogor (ID): Institut Pertanian Bogor.
- Hosmer DW, Lemeshow S. 2000. *Applied Logistic Regression*. Second Edition. Canada (USA): John Wiley & Sons, Inc.
- Johnson RA, Wichern DW. 2007. *Applied Multivariate Statistical Analysis*. Edisi ke-6. New Jersey: Prentice Hall.
- [PERKENI] Perkumpulan Endokrinologi Indonesia. 2015. Pengelolaan dan Pencegahan Diabetes Melitus Tipe 2 di Indonesia. Jakarta (ID): PERKENI.
- Puteri NA. 2018. Penentuan panjang gelombang lampu terbaik pada alat pemantau kadar glukosa darah non-invasif [skripsi]. Bogor (ID): Institut Pertanian Bogor.
- Sari DRP. 2019. Klasifikasi kadar glukosa darah menggunakan metode regresi logistik ordinal pada data tidak seimbang [tesis]. Bogor (ID): Institut Pertanian Bogor.
- Sari PP, Herdiani ET, Sunusi N. 2021. Pendeteksian pencilan menggunakan algoritma minimum vector variance dengan fungsi depth dan jarak mahalanobis. *J. Matematika, Statistika, dan Komputasi*. 17(3): 418-427
- Satrianto AGM. 2020. Penerapan Peringkasan Luas pada Pemulusan Kernel untuk Data Hasil Pengukuran Alat Pemantau Glukosa Darah Non-Invasif [Skripsi]. Bogor (ID): Institut Pertanian Bogor.
- Wibawati Y. 2021. Klasifikasi kadar glukosa darah keluaran alat non-invasif menggunakan support vector machine dan random forest [tesis]. Bogor (ID): Institut Pertanian Bogor.