



PERBANDINGAN METODE SMOTE, BORDERLINE-SMOTE, DAN ADASYN PADA KLASIFIKASI KADAR GULA DARAH NON-INVASIF DENGAN PERINGKASAN LUAS

ALWAN DAMAR FARRASTRA



**PROGRAM STUDI STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Perbandingan Metode SMOTE, Borderline-SMOTE, dan ADASYN pada Klasifikasi Data Kadar Gula Darah Non-Invasif dengan Peringkasan Luas” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juli 2025

Alwan Damar Farrastra
G14180075

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

ALWAN DAMAR FARRASTRA. Perbandingan Metode SMOTE, Borderline-SMOTE, dan ADASYN pada Klasifikasi Data Kadar Gula Darah Non-Invasif dengan Peringkasan Luas. Dibimbing oleh ERFIANI dan MOHAMMAD MASJKUR.

Diabetes melitus merupakan penyakit kelainan yang sering disebut sebagai *silent killer*. Prevalensi penderita pada tahun 2017 mencapai 8,8% atau 425 juta orang dan diprediksi pada tahun 2045 mencapai 9,9% atau sekitar 700 juta orang di dunia. Metode invasif menggunakan glukometer menjadi metode yang paling sering dijumpai, tapi karena caranya yang melukai jari membuat sedikit ketidaknyamanan dalam penggunaannya. Tim non-invasif *biomarking* IPB sedang melakukan pengembangan metode non-invasif dengan menggunakan alat spektroskopi inframerah yang menghasilkan *output* data spektrum residu intensitas. Data residu intensitas cahaya pada setiap puncak pada amatan dan tiap ulangan dapat terbaca berbeda sehingga memiliki banyak amatan tidak sama. Perlu dilakukannya penanganan penyeragaman data, salah satunya peringkasan luas puncak. Hubungan antara nilai kadar glukosa darah keluaran alat invasif dengan nilai residu intensitas keluaran alat non-invasif dapat dianalisis menggunakan metode klasifikasi. Tujuan penggunaan metode klasifikasi adalah untuk melihat seberapa tepat alat pengukur darah non-invasif dalam mengelompokkan nilai residu intensitas berdasarkan kelas kadar glukosa darah. Metode klasifikasi sendiri terkadang mendapatkan kendala berupa data yang tidak seimbang. Perlu dilakukan penanganan data tidak seimbang, salah satunya dengan *Synthetic Minority Oversampling Technique* (SMOTE), *Borderline-Synthetic Minority Oversampling Technique* (Borderline-SMOTE), dan *Adaptive Synthetic* (ADASYN). Klasifikasi menggunakan model *random forest* dengan ketiga metode dilakukan untuk menentukan hasil klasifikasi terbaik. Penelitian bertujuan untuk mengidentifikasi metode penanganan data tidak seimbang terbaik menggunakan model klasifikasi *random forest*. Evaluasi hasil penanganan SMOTE, Borderline-SMOTE, dan ADASYN menunjukkan penanganan Borderline-SMOTE menghasilkan akurasi paling tinggi dengan persentase 69,57%. Penanganan SMOTE dianggap memiliki persentase terkecil dengan 60,87% dan ADASYN menghasilkan akurasi 65,22%.

Kata kunci: alat ukur non-invasif, *adasyn*, *borderline-smote*, *random forest*, *smote*.



ABSTRACT

ALWAN DAMAR FARRASTRA. Comparison of SMOTE, Borderline-SMOTE, and ADASYN Methods on Classification of Non-Invasive Blood Glucose Level Data with Graph Area Summary. Supervised by ERFIANI and MOHAMMAD MASJKUR.

Diabetes mellitus is a disorder that is often referred to as the silent killer. Prevalence of sufferers in 2017 reaching 8.8% or 425 million people and predicted in 2045 to reach 9.9% or around 700 million people in the world. The invasive method of using a glucometer is the most common method, but because of the way it hurts the fingers, it makes it a little uncomfortable to use. IPB's non-invasive biomarking team is developing a non-invasive method using infrared spectroscopy that produces intensity residual spectrum data output. The residual data of light intensity at each peak at each observation and each replicate can be read differently so that there are many different observations. It is necessary to handle data uniformity, one of which is the summary of the graph area. The relationship between the blood glucose level value of the output of the invasive device and the residual value of the intensity of the output of the non-invasive device can be analyzed using the classification method. The purpose of using the classification method is to see how accurate non-invasive blood measuring devices are in classifying intensity residue values based on blood glucose level class. The classification method itself sometimes gets obstacles in the form of unbalanced data. It is necessary to handle unbalanced data, one of which is the Synthetic Minority Oversampling Technique (SMOTE), Borderline-Synthetic Minority Oversampling Technique (Borderline-SMOTE), and Adaptive Synthetic (ADASYN). Classification uses a random forest model with all three methods to determine the best classification results. The study aims to identify the best method of handling unbalanced data using a random forest classification model. The evaluation of the results of SMOTE, Borderline-SMOTE, and ADASYN handling showed that the Borderline-SMOTE handling produced the highest accuracy with a percentage of 69.57%. SMOTE handling is considered to have the smallest percentage with 60.87% and ADASYN produces an accuracy of 65.22%.

Keywords: adasyn, borderline-smote, non-invasive device, smote, random forest.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025¹
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



PERBANDINGAN METODE SMOTE, BORDERLINE-SMOTE, DAN ADASYN PADA KLASIFIKASI KADAR GULA DARAH NON-INVASIF DENGAN PERINGKASAN LUAS

ALWAN DAMAR FARRASTRA

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
Gerry Alfa Dito, S.Si., M.Si



Judul Skripsi : Perbandingan Metode SMOTE, Borderline-SMOTE, dan
ADASYN pada Klasifikasi Data Kadar Gula Darah Non-Invasif
dengan Peringkasan Luas

Nama : Alwan Damar Farrastra
NIM : G14180075

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:
Dr. Ir. Erfiani, M.Si.

Pembimbing 2:
Ir. Mohammad Masjkur, MS.



Diketahui oleh

Ketua Program Studi:
Dr. Bagus Sartono, M.Si.
NIP. 197804112005011002

Tanggal Ujian:
2 Januari 2025

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga penulis dapat menyelesaikan karya ilmiah dengan judul Perbandingan Metode SMOTE, Borderline-SMOTE, dan ADASYN pada Klasifikasi Data Kadar Gula Darah Non-Invasif dengan Peringkasan Luas.

Penulisan karya ilmiah ini tidak terlepas dari berbagai pihak yang telah mendukung dan membantu sehingga karya ilmiah ini dapat terselesaikan dengan baik. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Ibu Dr. Ir. Erfiani, M.Si., dan Bapak Ir. Mohammad Masjkur, MS., selaku komisi pembimbing yang telah memberikan bimbingan, arahan, saran, dan dukungan selama penyusunan karya ilmiah ini.
2. Dosen penguji luar, Bapak Gerry Alfa Dito, S.Si., M.Si. yang telah memberikan saran dan masukan dalam penulisan karya ilmiah ini.
3. Seluruh dosen dan pegawai administrasi Departemen Statistika yang telah memberikan ilmu yang bermanfaat dan menunjang segala kebutuhan administrasi penulis selama perkuliahan dan penyusunan karya ilmiah ini.
4. Kedua orang tua, teteh, kedua adik, dan keluarga besar penulis yang selalu memberikan doa, dukungan, dan kasih sayang kepada penulis.
5. Teman-teman Statistika 55 dan sahabat-sahabat penulis yang senantiasanya mendukung dan kebersamaian dalam penyusunan karya ilmiah ini.
6. Seluruh pihak yang telah membantu penulis dalam penyusunan karya ilmiah ini yang tidak dapat penulis sebutkan satu persatu.

Penulis berharap karya ilmiah ini dapat bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Juli 2025

Alwan Damar Farrastra



DAFTAR ISI

DAFTAR ISI	x
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xi
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Tujuan	2
TINJAUAN PUSTAKA	3
2.1 Alat Keluaran Non-Invasif	3
2.2 Imbalanced Data	3
2.3 Synthetic Minority Oversampling Technique (SMOTE)	4
2.4 Borderline-Synthetic Minority Oversampling Technique (Borderline-SMOTE)	4
2.5 Adaptive Synthetic (ADASYN)	4
2.6 Random Forest	5
2.7 Ukuran Kebaikan Model	5
III METODE	7
3.1 Data	7
3.2 Prosedur Analisis Data	7
IV HASIL DAN PEMBAHASAN.....	10
4.1 Eksplorasi Data	10
4.2 Penanganan Data Tak Seimbang	12
4.3 Evaluasi dengan <i>Random Forest</i>	14
4.4 Evaluasi Metode Penanganan Terbaik pada Penelitian Sebelumnya	15
V SIMPULAN DAN SARAN	17
5.1 Simpulan	17
5.2 Saran	17
DAFTAR PUSTAKA.....	18
LAMPIRAN	20
RIWAYAT HIDUP	25



DAFTAR TABEL

1	Matriks konfusi untuk tiga kelas	6
2	Perhitungan kebaikan model dengan <i>confusion matrix</i> untuk klasifikasi 3 kelas	6
3	Deskripsi kadar glukosa darah keluaran dari metode invasif	10
4	Proporsi kelas kadar glukosa darah keluaran alat metode invasif	11
5	Proporsi kelas kadar glukosa darah pada data <i>train</i> dengan SMOTE	13
6	Proporsi kelas kadar darah pada data <i>train</i> dengan Borderline-SMOTE	13
7	Proporsi kelas kadar glukosa darah pada data <i>Train</i> dengan ADASYN	14
8	Perbandingan evaluasi dengan <i>random forest</i>	15
9	Metode terbaik pada penelitian sebelumnya	16

DAFTAR GAMBAR

10	Ilustrasi metode alat ukur glukosa darah non-invasif	3
11	Ilustrasi perhitungan luas puncak	8
12	Bagan alir penelitian	9
13	Diagram <i>box-plot</i> kadar gula darah responden	10
14	Histogram dari kadar gula darah responden	11
15	Grafik keluaran alat metode non-invasif pada salah satu responden secara acak	12

DAFTAR LAMPIRAN

16	Lampiran 1 Data alat kadar glukosa darah dengan metode non-invasif	21
17	Lampiran 2 Data sebelum dilakukan peringkasan luas	22
18	Lampiran 3 Data setelah dilakukan peringkasan luas	23
19	Lampiran 4 Data keluaran metode invasif	24