



**KOMPRAISI DATA IMBLANCE DATA MENGGUNAKAN SMOTE-RBO,
RN-SMOTE, SMOTE-ENN, SMOTE, DBSCAN-SMOTE DENGAN
ATRIBUT STROKE**

LAPORAN TUGAS AHKIR

**IBNU HERMANSYAH
41521010014**

**UNIVERSITAS
MERCU BUANA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**KOMPRASI DATA IMBLANCE DATA MENGGUNAKAN SMOTE-RBO,
RN-SMOTE, SMOTE-ENN, SMOTE, DBSCAN-SMOTE DENGAN
ATRIBUT STROKE**

LAPORAN TUGAS AHKIR

**IBNU HERMANSYAH
41521010014**

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**UNIVERSITAS
MERCU BUANA
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : IBNU HERMANSYAH
NIM : 41521010014
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Komprasi Data Imblance Data Menggunakan
Smote-RBO, RN-Smote, Smote-ENN, Smote,
DBSCAN-Smote Dengan Atribut Stroke

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Proposal Penelitian saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Jakarta, 4 Agustus 2025



UNIVERSITAS
MERCU BUANA

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : IBNU HERMANSYAH
NIM : 41521010014
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Komprasi Data Imblance Data Menggunakan Smote-RBO, RN-Smote, Smote-ENN, Smote, DBSCAN-Smote Dengan Atribut Stroke

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Wawan Gunawan, S.Kom.,
M.Kom
NIDN : 0424108104
Ketua Penguji : Inna Sabily Karima, S.Kom.,
M.Kom
NIDN : 0324018902
Penguji 1 : Lukman Hakim, ST, M.Kom
NIDN : 0327107701
Penguji 2 : Prastika Indriyanti, S.Kom., MCS
NIDN : 0312089401



Jakarta, 4 Agustus 2025

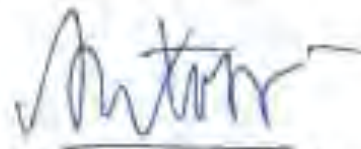
Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jekonowo, S.Si., MTI
NIDN : 0320037002



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan proposal penelitian yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa proposal penelitian ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Bapak Wawan Gunawan, S.Kom., M.Kom. selaku dosen pembimbing MPTI yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana..
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 29 Juli 2025

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : IBNU HERMANSYAH
NIM : 41521010014
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Komprasi Data Imblance Data Menggunakan Smote-RBO, RN-Smote, Smote-ENN, Smote, DBSCAN-Smote Dengan Atribut Stroke

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 4 Agustus 2025

Yang menyatakan,



Ibnu Hermansyah

ABSTRAK

Nama : IBNU HERMANSYAH
NIM : 41521010014
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Komprasi Data Imblance Data Menggunakan Smote-RBO, RN-Smote, Smote-ENN, Smote, DBSCAN-Smote Dengan Atribut Stroke
Dosen Pembimbing : Wawan Gunawan, S.Kom., M.Kom

Pada penyakit Stroke atau dikenal dengan Cerebrovascular Accident (CVA), merupakan penyakit serius yang menjadi penyebab utama kematian dan kecacatan di dunia. Prevalensi stroke terus meningkat, terutama di negara berkembang, termasuk Indonesia. Perkembangan teknologi medis memungkinkan penggunaan machine learning untuk mendukung deteksi dini dan prediksi stroke, yang dapat membantu intervensi lebih cepat. Namun, ketidakseimbangan kelas dalam dataset dapat memengaruhi akurasi prediksi dan mengarah pada bias terhadap kelas mayoritas. Penelitian ini membandingkan kinerja algoritma yang lainnya dalam klasifikasi penyakit stroke dengan menerapkan Synthetic Minority Oversampling Technique (SMOTE) untuk menangani ketidakseimbangan data. Dari hasil Riskesdas Kemenkes RI 2018, di Indonesia memiliki jumlah penderita stroke 713.783. Di provinsi NTB mengalami sebuah peningkatan penyakit stroke dari tahun 2013 samapi pada tahun 2018 pada pervalensi yang memiliki jumlah 4,5% menjadi 8%. Machine learning pendekatan ini dapat memprediksi suatu penyakit. Dengan adanya pendekatan pada Machine Learning ini, diteukannya imblance class pada dataset yang di gunakan. Hal ini membuat imblance classs terjadi karena kelas majority jauh lebih tinggi dibandingkan class minority Hasil penelitian menunjukkan bahwa kombinasi SMOTE dan algoritma meningkatkan akurasi prediksi serta performa model dibandingkan metode konvensional. Implementasi teknik ini dapat menjadi solusi efektif dalam meningkatkan deteksi dini stroke dan memberikan kontribusi bagi dunia medis untuk pencegahan penyakit yang lebih baik.

Kata kunci: Stroke, CVA, kecacatan, indonesia, Prediksi Stroke, Machine Learning, Teknologi medis, Data tidak Simbang, privalensi, Pencegahan penyakit, Oversampling, SMOTE, Klasifikasi, Data Mining, Prediksi Kesehatan, Indonesia.

ABSTRACT

Nama : IBNU HERMANSYAH
NIM : 41521010014
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Komprasi Data Imblance Data Menggunakan
Smote-RBO, RN-Smote, Smote-ENN, Smote,
DBSCAN-Smote Dengan Atribut Stroke
Dosen Pembimbing : Wawan Gunawan, S.Kom., M.Kom

Stroke, also known as Cerebrovascular Accident (CVA), is a serious disease that is a leading cause of death and disability worldwide. The prevalence of stroke continues to increase, especially in developing countries, including Indonesia. Advances in medical technology enable the use of machine learning to support early detection and stroke prediction, which can facilitate faster intervention. However, class imbalance in the dataset can affect prediction accuracy and lead to bias towards the majority class. This study compares the performance of other algorithms in stroke classification by applying the Synthetic Minority Oversampling Technique (SMOTE) to address data imbalance. Based on the 2018 Riskesdas (Basic Health Research) of the Indonesian Ministry of Health, Indonesia has 713,783 stroke sufferers. West Nusa Tenggara (NTB) province experienced an increase in stroke cases from 2013 to 2018, with a prevalence of 4.5% to 8%. This machine learning approach can predict a disease. With this Machine Learning approach, class imbalance was found in the dataset used. This creates imbalanced classes because the majority class is much higher than the minority class. The results show that the combination of SMOTE and the algorithm improves prediction accuracy and model performance compared to conventional methods. Implementing this technique could be an effective solution for improving early stroke detection and contributing to better disease prevention in the medical world.

Keywords: Stroke, CVA, disability, Indonesia, Stroke Prediction, Machine Learning, Medical technology, Imbalanced data, privacy, Disease prevention, Oversampling, SMOTE, Classification, Data Mining, Health Prediction, Indonesia.

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN.....	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	5
2.1 Teori Pendukung.....	5
2.2 Penelitian Terdahulu.....	8
BAB III METODE PENELITIAN	18
2.3 Teori Utama	18
3.1 Jenis Penelitian.....	19
3.2 Tahapan Penelitian.....	20
BAB IV HASIL YANG DIHARAPKAN.....	23
4.1. Dataset.....	23
4.2 Pre-processing.....	27
4.3 Metode smote.....	29
4.4 Smote Gabungan	35
4.5 Algoritma	37
4.6 Hasil evaluasi	37
BAB V KESIMPULAN DAN SARAN.....	41
5.1 Kesimpulan	41

5.2	Saran	41
DAFTAR PUSTAKA.....		43
LAMPIRAN.....		46



DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	9
Tabel 4.1 Dataset Awal	24
Tabel 4.2 Dataset Processing	25
Tabel 4.3 DBSCAN	31
Tabel 4.4 Acuraccy DBSCAN	31
Tabel 4.5 RN SMOTE	32
Tabel 4.6 Acuraccy RN SMOTE	32
Tabel 4.7 SMOTE ENN	33
Tabel 4.8 Acuraccy SMOTE ENN	33
Tabel 4.9 SMOTE	34
Tabel 5.0 Acuraccy SMOTE	34
Tabel 5.1 SMOTE RBO	35
Tabel 5.2 Acuraccy SMOTE RBO	35
Tabel 5.3 Metode Smote Gabungan	36



DAFTAR GAMBAR

Gambar 2.1 Logo Google Colab	8
Gambar 3.1 Tahapan Penelitian	20
Gambar 4.1 Tahapan Kategori	28
Gambar 4.2 Data Plot.....	29
Gambar 4.3 Diagram Sex.....	29
Gambar 4.4 Distribusi Kategori Sex	30
Gambar 4.5 Heat Map Model	39
Gambar 4.6 Performa Model	39
Gambar 4.7 Model Machine Learning.....	40
Gambar 4.8 Heat Map.....	41



DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi	46
Lampiran 2 Curriculum Vitae	47
Lampiran 3 Surat Pernyataan HAKI.....	48
Lampiran 4 Hasil Cek Turnitin	49
Lampiran 5 Hasil Persetujuan	50
Lampiran 6 Similarity check.....	51
Lampiran 7 Lembar revisi (Penguji 1).....	52
Lampiran 8 Lembar revisi (Penguji 2).....	53
Lampiran 9 BNSP	54
Lampiran 10 Surat Pernyataan Pengalihan Hak Cipta.....	55

