SKRIPSI

IMPLEMENTASI ALGORITMA *RANDOM FOREST* DALAM PREDIKSI *STROKE* DENGAN PENGGUNAAN *SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE*

JOHANES GERALD

NIM. 2010511106

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

2023

SKRIPSI

IMPLEMENTASI ALGORITMA *RANDOM FOREST* DALAM PREDIKSI *STROKE* DENGAN PENGGUNAAN *SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE*

JOHANES GERALD

NIM. 2010511106

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

2023

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua seumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Johanes Gerald
NIM : 2010511106
Tanggal : 29 November 2023

Bilamina di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 30 November 2023

Yang Menyatakan,

Johanes Gerald

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Johanes Gerald
NIM : 2010511106
Fakultas : Ilmu Komputer
Program Studi : S1 Informatika

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right)* atas karya ilmiah saya yang berjudul:

IMPLEMENTASI ALGORITMA *RANDOM FOREST* DALAM PREDIKSI *STROKE* DENGAN PENGGUNAAN *SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Dibuat di : Jakarta
Pada Tanggal : 30 November 2023

Yang menyatakan,

Johanes Gerald

LEMBAR PENGESAHAN

Tugas akhir ini diajukan oleh:

Nama	:	Johanes Gerald
NIM	:	2010511106
Program Studi	:	S1-Informatika
Judul	:	IMPLEMENTASI ALGORITMA *RANDOM FOREST* DALAM PREDIKSI *STROKE* DENGAN PENGGUNAAN *SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE*

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Dr. Ermatita, M.Kom
Penguji 1

Ika Nurlaili Isnainiyah, S.Kom., M.Sc.
Penguji 2

Dr. Widya Cholil, M.I.T
Pembimbing

Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM
Dekan

Dr. Widya Cholil, M.I.T
Kepala Program Studi

Ditetapkan di : Jakarta
Tanggal Ujian : 10 Januari 2024

IMPLEMENTASI ALGORITMA *RANDOM FOREST* DALAM PREDIKSI *STROKE* DENGAN PENGGUNAAN *SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE*

Johanes Gerald

ABSTRAK

Stroke adalah salah satu masalah kesehatan global yang signifikan dan menjadi penyebab utama kecacatan serta kematian di seluruh dunia. Di Indonesia, penyakit *stroke* telah menjadi salah satu penyumbang penyakit paling mematikan. Menurut data pada profil kesehatan Indonesia tahun 2020, *stroke* menempati posisi ketiga dengan jumlah kasus sebanyak 1.789.261. Tujuan penelitian ini adalah untuk mengidentifikasi pasien yang berisiko tinggi terkena *stroke*. Algoritma yang digunakan adalah *Random Forest Classifier* dengan menggunakan *Synthetic Minority Oversampling Technique* untuk menyeimbangkan kelas data. Pada metode *Random Forest* menggunakan *Synthetic Minority Oversampling Technique* didapatkan hasil akurasi sebesar 95,61%, nilai presisi sebesar 93,66%, nilai recall sebesar 97,85%, dan nilai f1-score sebesar 95,71%. Sedangkan untuk model *Random Forest* didapatkan hasil akurasi sebesar 90,15%, nilai presisi sebesar 90,5%, nilai *recall* sebesar 90,15%, dan nilai *f1-score* sebesar 90,32%. Karena terjadi *imbalance class* maka algoritma *Random Forest* saja tidak tepat digunakan tanpa melakukan *resampling*. Sehingga *Random Forest* – SMOTE dapat digunakan sebagai salah satu algoritma untuk memprediksi *stroke*.

Kata kunci: Klasifikasi, *Stroke*, *Random Forest*, SMOTE.

IMPLEMENTATION OF RANDOM FOREST ALGORITHM IN STROKE PREDICTION USING SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE

Johanes Gerald

ABSTRACT

Stroke is one of the significant global health issues, being a leading cause of disability and death worldwide. In Indonesia, stroke has emerged as one of the most fatal diseases. According to the 2020 Indonesia health profile data, stroke ranks third with a total of 1,789,261 reported cases. The aim of this research is to identify patients at high risk of stroke. The algorithm employed is the Random Forest Classifier utilizing the Synthetic Minority Oversampling Technique to balance the class data. In the Random Forest method using the Synthetic Minority Oversampling Technique, the results showed an accuracy of 95.61%, precision of 93.66%, recall of 97.85%, and an f1-score of 95.71%. Meanwhile, for the Random Forest model, the accuracy was 90.15%, precision was 90.5%, recall was 90.15%, and the f1-score was 90.32%. Due to class imbalance, using the Random Forest algorithm alone is not suitable without resampling. Therefore, Random Forest – SMOTE can be utilized as one of the algorithms to predict strokes.

Keywords**:** *Classification, Stroke, Random Forest,* SMOTE.

KATA PENGANTAR

Puji Syukur penulis panjatkan kehadirat Tuhan Yesus Kristus, yang senantiasa memberikan petunjuk, kelancaran serta kemudahan kepada penulis dalam menyelesaikan skripsi ini, serta kasih anugerahNya yang tidak pernah berkesudahan. Penulisan skripsi ini bertujuan untuk memenuhi salah satu prasyarat untuk memperoleh gelar Sarjana Komputer, Jurusan Informatika.

Dalam penulisan skripsi ini, penulis mendapat banyak dukungan serta bantuan dari berbagai pihak, baik berupa materi, spiritual, dan informasi. Pada kesempatan kali ini, penulis mengucapkan terimakasih kepada:

1. Tuhan Yesus Kristus.
2. Kedua orang tua, cici serta keluarga tercinta yang selalu mendoakan serta mendukung penulis sehingga dapat menyelesaikan skripsi ini.
3. Ibu Dr. Widya Cholil, M.I.T, selaku dosen pembimbing dan Kaprodi Informatika yang berjasa dan memberikan bimbingan hingga terselesaikannya skripsi ini
4. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM selaku Dekan Fakultas Ilmu Komputer.
5. Ibu Dr. Ermatita, M.Kom dan Ibu Ika Nurlaili Isnainiyah, S.Kom., M.Sc selaku dosen penguji.
6. Seluruh jajaran Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta yang telah membantu dalam perizinan dan administrasi.
7. Teman – teman Informatika yang berjuang bersama selama perkuliahan, memberikan semangat dan dorongan untuk dapat menyelesaikan kuliah dan skripsi ini.
8. Kakak tingkat penulis yang seringkali menjadi tempat berdiskusi yaitu Bang Ojan, Bang Kipau, Bang Radit.
9. Prambanan Crew yaitu Gilbert, Tito, Sarah, Nida, Billy, Wildan dan sahabat – sahabat perjuangan yang selalu membantu dan memberikan masukan serta doa kepada penulis agar terselesaikannya skripsi ini dengan baik.
10. Terima kasih banyak kepada semua pihak atas bantuan, dukungan, semangat, dan doa yang tidak dapat penulis sebutkan satu persatu

11. Johanes Gerald, *last but not least*, ya! diri saya sendiri. Apresiasi sebesar-besarnya karena telah bertanggung jawab untuk menyelesaikan apa yang telah dimulai. Terima kasih karena terus berusaha dan tidak menyerah, serta senantiasa menikmati setiap prosesnya yang bisa dibilang tidak mudah. Terima kasih sudah bertahan.

Penyusun menyadari bahwa penyusunan skripsi ini masih jauh dari sempurna, oleh karena itu penyusun mengharapkan kritik dan saran yang bersifat membangun untuk kesempurnaan skripsi ini. Akhir kata penyusun mengharapkan semoga skripsi ini dapat bermanfaat bagi semua pihak.

Jakarta, 10 Januari 2024

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN