



**PEPENGARUH SMOTE TERHADAP AKURASI KLASIFIKASI DATA
AUTISM SPECTRUM DISORDER PADA ORANG DEWASA DENGAN
ALGORITMA RANDOM FOREST: STUDI KASUS PADA DATASET TIDAK
SEIMBANG**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

MUHAMMAD HAFIZH MUSYAFFA

NIM 2111016310001

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

MEI 2025



**PEPENGARUH SMOTE TERHADAP AKURASI KLASIFIKASI DATA
AUTISM SPECTRUM DISORDER PADA ORANG DEWASA DENGAN
ALGORITMA RANDOM FOREST: STUDI KASUS PADA DATASET TIDAK
SEIMBANG**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

MUHAMMAD HAFIZH MUSYAFFA

NIM 2111016310001

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

MEI 2025

SKRIPSI

PENGARUH SMOTE TERHADAP AKURASI KLASIFIKASI DATA AUTISM SPECTRUM DISORDER PADA ORANG DEWASA DENGAN ALGORITMA RANDOM FOREST: STUDI KASUS PADA DATASET TIDAK SEIMBANG

Oleh:

MUHAMMAD HAFIZH MUSYAFFA
NIM. 2111016310001

Telah dipertahankan di depan Dosen Penguji pada tanggal 16 Mei 2025.

Susunan Dosen Penguji:

Pembimbing I



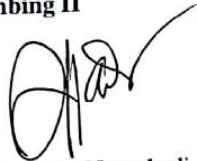
Triando Hamonangan Saragih, S.Kom., M.Kom.
NIP. 199308242019031012

Dosen Penguji I



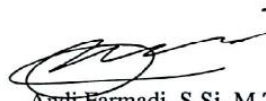
Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003

Pembimbing II



Dodon Turianto Nugrahadi, S.Kom, M.Eng
NIP. 198001122009121002

Dosen Penguji II



Andi Farmadi, S.Si, M.T
NIP. 197307252008011006

Banjarbaru, 16 Mei 2025

Koordinator Program Studi Ilmu Komputer



Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam jurnal ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 16 Mei 2025



Muhammad Hafizh Musyaffa

NIM.2111016310001

ABSTRAK

PEPENGARUH SMOTE TERHADAP AKURASI KLASIFIKASI DATA AUTISM SPECTRUM DISORDER PADA ORANG DEWASA DENGAN ALGORITMA RANDOM FOREST: STUDI KASUS PADA DATASET TIDAK SEIMBANG

(Oleh : Muhammad Hafizh Musyaffa , Triando Hamonangan Saragih , S.Kom., M.Kom., Dodon Turianto Nugrahadi S.Kom., M.Eng., 2025; halaman)

Autism Spectrum Disorder (ASD), yang awalnya dideskripsikan oleh Leo Kanner pada tahun 1943, merupakan kondisi perkembangan kompleks yang bermanifestasi melalui tantangan sosial, emosional, dan perilaku, yang sering kali meliputi keterlambatan bicara dan kesulitan dalam interaksi interpersonal. Meskipun ada kemajuan yang signifikan dalam kriteria diagnostik selama bertahun-tahun, diagnosis ASD yang akurat pada orang dewasa tetap menjadi tantangan karena terbatasnya akses ke kumpulan data yang komprehensif dan kendala metodologis yang melekat. Dataset Autism Screening Adult yang digunakan dalam penelitian ini mencontohkan masalah-masalah ini, karena mengandung nilai yang hilang dan menunjukkan ketidakseimbangan kelas yang nyata, yang keduanya dapat mempengaruhi kinerja model. Untuk mengatasi tantangan ini, kami mengusulkan kerangka kerja yang mengintegrasikan klasifikasi Random Forest dengan imputasi MissForest dan Synthetic Minority Over-sampling Technique (SMOTE). MissForest secara efektif mengimputasi data yang hilang dengan menggunakan pendekatan hutan acak berulang yang mempertahankan struktur dasar data tanpa bergantung pada asumsi parametrik yang ketat. Sementara itu, SMOTE menghasilkan sampel sintetis untuk kelas minoritas, sehingga menyeimbangkan kumpulan data dan mengurangi bias prediksi. Evaluasi eksperimental melalui 10-Fold Cross Validation menunjukkan bahwa penerapan SMOTE secara signifikan meningkatkan kinerja model. Secara khusus, akurasi keseluruhan meningkat dari 70,17% menjadi 79,32%, dan AUC-ROC meningkat dari 47,13% menjadi 85,84%, yang mengindikasikan peningkatan yang kuat dalam kemampuan model untuk membedakan antara kasus positif dan negatif. Hasil ini menggarisbawahi pentingnya mengatasi ketidakseimbangan data dan nilai yang hilang dalam pemodelan prediktif untuk ASD.

Kata – kunci : Autism Spectrum Disorder, Imbalanced Data, Missing Values, MissForest, Random Forest, SMOTE

ABSTRACT

PEPENGARUH SMOTE TERHADAP AKURASI KLASIFIKASI DATA AUTISM SPECTRUM DISORDER PADA ORANG DEWASA DENGAN ALGORITMA RANDOM FOREST: STUDI KASUS PADA DATASET TIDAK SEIMBANG

(By : Muhammad Hafizh Musyaffa , Triando Hamonangan Saragih , S.Kom., M.Kom., Dodon Turianto Nugrahadi S.Kom., M.Eng., 2025; halaman)

Autism Spectrum Disorder (ASD), originally described by Leo Kanner in 1943, is a complex developmental condition that manifests through social, emotional, and behavioral challenges, often including speech delays and difficulties in interpersonal interactions. Despite significant advancements in diagnostic criteria over the years, accurate diagnosis of ASD in adults remains challenging due to limited access to comprehensive datasets and inherent methodological constraints. The Autism Screening Adult dataset used in this study exemplifies these issues, as it contains missing values and exhibits a marked class imbalance, both of which can adversely affect model performance. To address these challenges, we proposed a framework that integrates Random Forest classification with MissForest imputation and the Synthetic Minority Over-sampling Technique (SMOTE). MissForest effectively imputes missing data by employing an iterative random forest approach that preserves the underlying structure of the data without relying on strict parametric assumptions. Meanwhile, SMOTE generates synthetic samples for the minority class, thereby balancing the dataset and reducing prediction bias. Experimental evaluation through 10-Fold Cross Validation demonstrated that the application of SMOTE significantly enhanced model performance. Notably, the overall accuracy improved from 70.17% to 79.32%, and the AUC-ROC increased from 47.13% to 85.84%, indicating a robust improvement in the model's ability to distinguish between positive and negative cases. These results underscore the critical importance of addressing data imbalance and missing values in predictive modeling for ASD. The promising outcomes of this study provide a solid foundation for developing more reliable diagnostic tools for adult ASD, and future research may further refine feature selection and incorporate additional data sources to optimize performance even further.

Keywords : Autism Spectrum Disorder, Imbalanced Data, Missing Values, MissForest, Random Forest, SMOTE

KATA PENGANTAR

Puji syukur penulis panjatkan ke Tuhan kita Yang Maha Esa karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan jurnal yang berjudul *“Effectiveness of SMOTE in Enhancing Adult Autism Spectrum Disorder Diagnosis Predictive Performance With Missforest Imputation And Random Forest”* untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada semua pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan jurnal ini, adapun yang dimaksud adalah sebagai berikut:

1. Diri saya sendiri yang tidak pernah patah semangat walaupun banyak menemui kesulitan baik disebabkan oleh diri sendiri maupun hal lain.
2. Keluarga besar yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian jurnal ini.
3. Bapak Triando Hamonangan Saragih, S.Kom, M.Kom selaku dosen pembimbing utama yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
4. Bapak Dodon Turianto Nugrahadi, S.Kom, M.Eng. selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
5. Ibu Dwi Kartini, S.Kom., M.Kom. selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau jurnal ini dapat diselesaikan.
6. Seluruh Dosen dan staff Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
7. Rafi, Ryan, Iwan, Fikeri, Zaki, Adam, Naufal, dan Sugi yang selalu mendukung dalam berbagai hal.
8. Teman-teman dan sahabat-sahabat keluarga Ilmu Komputer yang memberikan dukungan dan selalu mengingatkan serta mendoakan dalam proses mengerjakan jurnal.
9. Serta semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian jurnal ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya dibidang informatika serta mendapat keridhaan Allah SWT.

Banjarbaru, 16 Mei 2025

A handwritten signature in black ink, consisting of stylized letters 'H' and 'M' followed by a horizontal line.

Muhammad Hafizh Musyaffa