



**THE EFFECT OF SMOTE-TOMEK ON THE CLASSIFICATION
OF CHRONIC DISEASES BASED ON HEALTH AND
LIFESTYLE DATA**

Skripsi

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

Oleh

Muhammad Adika Riswanda

NIM 2011016210025

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

Mei 2026



**THE EFFECT OF SMOTE-TOMEK ON THE CLASSIFICATION OF
CHRONIC DISEASES BASED ON HEALTH AND LIFESTYLE DATA**

Skripsi

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

**Oleh
Muhammad Adika Riswanda
NIM 2011016210025**

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

Mei 2026

SKRIPSI

THE EFFECT OF SMOTE-TOMEK ON THE CLASSIFICATION OF CHRONIC DISEASES BASED ON HEALTH AND LIFESTYLE DATA

Oleh:

Muhammad Adika Riswanda

NIM 2011016210025

telah dipertahankan di depan Dosen Penguji pada tanggal 20 Mei 2026.

Susunan Penguji:

Pembimbing Utama



Friska Abadi, S.Kom., M.Kom.

NIP. 198809132023211010

Penguji

1.



Mohammad Reza Faisal, S.Si., S.T., M.T., Ph.D.

NIP. 197612202008121001

Pembimbing Pendamping



Muhammad Itqan Mazdadi, S.Kom., M.Kom.

NIP. 199006122019031013

2.



Rudy Herteno, S.Kom., M.Kom.

NIP. 198809252022031003

Banjarbaru, 21 Mei 2026

Koordinator Program Studi Ilmu Komputer



Dwi Kartini, S.Kom., M.Kom.

NIP. 198704212012122003

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas rahmat, karunia, serta pertolongan-Nya, sehingga penulis dapat menyelesaikan jurnal yang berjudul “*The Effect Of Smote-Tomek On The Classification Of Chronic Diseases Based On Health And Lifestyle Data*”.

Jurnal ini disusun sebagai salah satu syarat untuk menyelesaikan studi pada Program Studi Ilmu Komputer FMIPA Universitas Lambung Mangkurat. Dalam proses penyusunannya, penulis menyadari bahwa penyelesaian jurnal ini tidak lepas dari dukungan, bantuan, serta bimbingan dari berbagai pihak yang turut berperan selama penelitian berlangsung. Oleh karena itu, penulis menyampaikan rasa terima kasih dan penghargaan setinggi-tingginya kepada:

1. Allah SWT, yang senantiasa memberikan kesehatan, kekuatan, dan kemudahan sehingga jurnal ini dapat diselesaikan.
2. Diri penulis sendiri, atas keteguhan hati dan semangat yang terus terjaga dalam menghadapi setiap tantangan selama penyusunan jurnal.
3. Keluarga tercinta, yang selalu menjadi penyemangat utama melalui doa, dukungan moral, serta kasih sayang yang tidak pernah putus.
4. Bapak Friska Abadi, S.Kom., M.Kom., selaku dosen pembimbing utama yang telah memberikan arahan, bimbingan, dan masukan selama proses penyusunan jurnal ini.
5. Bapak Muhammad Itqan Mazdadi, S.Kom., M.Kom., selaku dosen pembimbing pendamping yang dengan penuh kesabaran turut membimbing serta memberikan saran demi penyempurnaan jurnal ini.
6. Ibu Dwi Kartini, S.Kom., M.Kom., selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM, atas dukungan dan kebijakan yang membantu kelancaran penyusunan jurnal.
7. Seluruh dosen serta staf Program Studi Ilmu Komputer FMIPA ULM, yang telah memberikan ilmu, bantuan, dan pelayanan akademik yang sangat bermanfaat bagi penulis.
8. Teman-teman serta sahabat Ilmu Komputer, yang telah memberikan dukungan,

peringat, dan doa selama proses penyusunan jurnal berlangsung.

9. Semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah membantu baik secara langsung maupun tidak langsung dalam proses penyusunan jurnal ini.

Akhir kata, penulis berharap jurnal ini dapat memberikan manfaat serta menambah wawasan bagi pembaca. Semoga Allah SWT senantiasa melimpahkan keberkahan dalam setiap langkah dan usaha kita.

Banjarbaru, 20 Mei 2026



Muhammad Adika Riswanda
NIM. 2011016210025

ABSTRAK

PENGARUH SMOTE-TOMEK TERHADAP KLASIFIKASI PENYAKIT KRONIS BERDASARKAN DATA KESEHATAN DAN GAYA HIDUP

(Oleh: Muhammad Adika Riswanda, Pembimbing: Friska Abadi, S.Kom., M.Kom., and Muhammad Itqan Mazdadi, S.Kom., M.Kom.; 2026; halaman)

Model pembelajaran mesin untuk prediksi penyakit kronis sering kali dilatih menggunakan dataset layanan kesehatan yang tidak seimbang, di mana kasus non-penyakit mendominasi. Kondisi ini dapat menyebabkan akurasi yang tampak tinggi namun menyesatkan, sekaligus gagal mengidentifikasi pasien dengan penyakit kronis, sehingga membatasi kegunaan klinisnya. Penelitian ini bertujuan untuk menganalisis dampak ketidakseimbangan kelas terhadap kinerja model dan mengevaluasi keefektifan teknik resampling SMOTE–Tomek dalam meningkatkan prediksi penyakit kronis. Penelitian ini memberikan bukti empiris bahwa akurasi saja tidak cukup untuk mengevaluasi model kesehatan dan menunjukkan bahwa prapemrosesan yang menyadari ketidakseimbangan sangat penting untuk deteksi penyakit kronis yang valid dan andal. Lima model klasifikasi, seperti Support Vector Machine, Random Forest, K-Nearest Neighbors, Gradient Boosting, dan XGBoost, dievaluasi pada dataset penyakit kronis berbasis gaya hidup di bawah dua kondisi: tanpa resampling dan dengan SMOTE–Tomek. Kinerja model dievaluasi menggunakan akurasi, presisi, recall, F1-score, dan AUC. Tanpa SMOTE–Tomek, semua model gagal mendeteksi kasus penyakit kronis, menghasilkan recall dan F1-score mendekati nol meskipun akurasi melebihi 80%. Setelah menerapkan SMOTE–Tomek, peningkatan yang signifikan diamati pada semua model, terutama pada recall dan AUC. Support Vector Machine (SVM) mencapai kinerja keseluruhan terbaik, dengan akurasi 92,9%, presisi 92%, recall 93,9%, skor F1 0,93, dan AUC 0,98. Temuan ini menegaskan bahwa penanganan ketidakseimbangan kelas merupakan prasyarat untuk prediksi penyakit kronis yang bermakna. Peningkatan yang konsisten pada recall dan AUC di seluruh model yang dievaluasi menegaskan bahwa peningkatan tersebut berasal dari peningkatan keterpisahan kelas, bukan inflasi metrik. Pendekatan yang diusulkan mendukung sistem skrining dini dan dukungan keputusan yang lebih andal dalam perawatan kesehatan preventif.

Kata Kunci: Ketidakseimbangan, SMOTE-Tomek, Penyakit Kronis, Pembelajaran Mesin, Analisis Kesehatan

ABSTRACT

THE EFFECT OF SMOTE-TOMEK ON THE CLASSIFICATION OF CHRONIC DISEASES BASED ON HEALTH AND LIFESTYLE DATA

(By: Muhammad Adika Riswanda, Supervisors: Friska Abadi, S.Kom., M.Kom., and Muhammad Itqan Mazdadi, S.Kom., M.Kom.; 2026; pages)

Machine learning models for chronic disease prediction are often trained on imbalanced healthcare datasets, where non-disease cases dominate. This condition can lead to misleadingly high accuracy while failing to identify patients with chronic diseases, limiting clinical usefulness. This study aims to analyze the impact of class imbalance on model performance and to evaluate the effectiveness of the SMOTE–Tomek resampling technique in improving chronic disease prediction. This research provides empirical evidence that accuracy alone is insufficient for evaluating healthcare models and demonstrates that imbalance-aware preprocessing is essential for valid and reliable chronic disease detection. Five classification models, such as Support Vector Machine, Random Forest, K-Nearest Neighbors, Gradient Boosting, and XGBoost, were evaluated on a lifestyle-based chronic disease dataset under two conditions: without resampling and with SMOTE–Tomek. Model performance was assessed using accuracy, precision, recall, F1-score, and AUC. Without SMOTE–Tomek, all models failed to detect chronic disease cases, producing near-zero recall and F1-scores despite accuracy exceeding 80%. After applying SMOTE–Tomek, substantial improvements were observed across all models, particularly in recall and AUC. Support Vector Machine achieved the best overall performance, with an accuracy of 92.9%, a precision of 92%, a recall of 93.9%, an F1-score of 0.93, and an AUC of 0.98. The findings confirm that handling class imbalance is a prerequisite for meaningful chronic disease prediction. The consistent increase in recall and AUC across all evaluated models confirms that the improvement stems from enhanced class separability rather than metric inflation. The proposed approach supports more reliable early screening and decision-support systems in preventive healthcare.

Keywords: *Imbalance, SMOTE-Tomek, Chronic Disease, Machine Learning, Healthcare Analytics*



KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS LAMBUNG MANGKURAT
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
JURUSAN ILMU KOMPUTER
PROGRAM STUDI ILMU KOMPUTER

Jalan Jenderal Ahmad Yani KM 36, Banjarbaru, Kalimantan Selatan 70714
Telp/Fax (0511) 4773 112 Laman : <http://fmipa.ulm.ac.id>

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi 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, 20 Mei 2026
Yang Menyatakan,

Muhammad Adika Riswanda
NIM. 2011016210025