

**EVALUATING MULTIPLE CLASSIFICATION
ALGORITHMS FOR DIABETES DIAGNOSIS AND
PREDICTION ON AN IMBALANCED DATASET**

JALUR SCIENTIST



disusun oleh

MUHAMMAD ZEIN MAULANA AKBAR

22.12.2302

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**EVALUATING MULTIPLE CLASSIFICATION
ALGORITHMS FOR DIABETES DIAGNOSIS AND
PREDICTION ON AN IMBALANCED DATASET**

JALUR NON REGULER - SCIENTIST

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

MUHAMMAD ZEIN MAULANA AKBAR

22.12.2302

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

JALUR NON REGULER - SCIENTIST

**EVALUATING MULTIPLE CLASSIFICATION
ALGORITHMS FOR DIABETES DIAGNOSIS AND
PREDICTION ON AN IMBALANCED DATASET**

yang disusun dan diajukan oleh

MUHAMMAD ZEIN MAULANA AKBAR

22.12.2302

telah disetujui oleh Dosen Pembimbing
pada tanggal 7 February 2026

Dosen Pembimbing,



Afrig Aminuddin, S.Kom., M.Eng., Ph.D
NIK. 190302351

HALAMAN PENGESAHAN

JALUR SCIENTIST

EVALUATING MULTIPLE CLASSIFICATION
ALGORITHMS FOR DIABETES DIAGNOSIS AND
PREDICTION ON AN IMBALANCED DATASET

yang disusun dan diajukan oleh
MUHAMMAD ZEIN MAULANA AKBAR

22.12.2302

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 February 2026

Nama Penguji

Susunan Dewan Penguji

Tanda Tangan

Anggit Dwi Hartanto, S.Kom., M.Kom.
NIK. 190302163

Wiwid Widayani, S.Kom., M.Kom.
NIK. 190302272

Afriz Aminuddin, S.Kom., M.Eng., Ph.D.
NIK. 190302351



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 25 February 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KONTRIBUSI KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : MUHAMMAD ZEIN MAULANA AKBAR
NIM : 22.12.2302

Menyatakan bahwa karya dengan judul berikut:

Evaluating Multiple Classification Algorithms For Diabetes Diagnosis And Prediction On An Imbalanced Dataset

Dosen Pembimbing: Afrig Aminuddin, S.Kom., M.Eng., Ph.D

1. Karya adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya merupakan gagasan, rumusan maupun penelitian yang orisinal dan SAYA memiliki KONTRIBUSI terhadap karya tersebut.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka atau Referensi pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 13 January 2026

Yang Menyatakan,



Muhammad Zein Maulana Akbar

HALAMAN PERSEMBAHAN

Puji syukur kami haturkan ke hadirat Allah SWT, rahmat dan keberkahan Engkau telah memberikan kami kekuatan, melimpahi kami dengan ilmu serta memperkenalkan kami dalam keimanan. Karunia serta keberkahan yang engkau limpahkan menjadikan karya tugas akhir non skripsi dengan skema scientist terselesaikan dengan baik. Sholawat dan salam selalu ter-limpahkan keharibaan Rasulullah Muhammad SAW. Kami mempersembahkan karya sayang sederhana ini kepada orang-orang yang sangat kami kasihi dan sayangi.

1. Kedua orangtua kami; Bpk. Sukma Haitari dan Ibu Siti Rahmah yang telah senantiasa mendoakan, membimbing dan memotivasi kami untuk terus semangat. Semoga ini menjadi langkah awal untuk membuat Ayah dan Ibu ridho, karena kami sadar selama ini belum bisa berbuat lebih.
2. Saudara dan teman-teman seperjuangan yang telah banyak membantu kami dalam hal apapun. Semoga Allah memberi balasan dan ganti yang lebih baik kepada kalian semua orang-orang baik.
3. Dosen pembimbing kami, Afrig Aminuddin, S.Kom., M.Eng., Ph.D yang telah memberikan arahan dan bimbingan dalam penyusunan karya ilmiah ini, semoga limpahan ilmu Bapak senantiasa bermanfaat bagi semua orang.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT, Tuhan Yang Maha Esa, atas segala limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan penelitian serta penyusunan Artikel Ilmiah ini dengan baik. Shalawat serta salam semoga senantiasa tercurah kepada junjungan kita, Nabi Muhammad SAW, beserta keluarga, sahabat, dan para pengikutnya. Artikel Ilmiah yang berjudul, "Evaluating Multiple Classification Algorithms for Diabetes Diagnosis and Prediction on an Imbalanced Dataset" ini, disusun sebagai bagian dari pemenuhan publikasi ilmiah di 2025 IEEE 9th International Conference on Software Engineering & Computer Systems (ICSECS). Penulis menyadari sepenuhnya bahwa dalam proses penyusunan karya tulis ini, berbagai hambatan dan tantangan telah dihadapi. Namun, berkat pertolongan-Nya serta bimbingan, dukungan, dan doa dari berbagai pihak, semua kesulitan tersebut dapat teratasi. Penyelesaian Artikel Ilmiah ini tidak akan terwujud tanpa uluran tangan dan kontribusi dari berbagai pihak, baik secara moril maupun materiil. Oleh karena itu, pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih yang tulus dan penghargaan setinggi-tingginya kepada semua pihak yang telah membantu, terutama kepada:

1. Kedua Orangtua ter-sayang, atas segala curahan kasih sayang, dukungan moril maupun materiil, serta doa yang tiada putus yang telah di berikan kepada penulis.
2. Bapak Afrig Aminuddin, S.Kom., M.Eng., Ph.D., selaku dosen pembimbing. Penulis mengucapkan terima kasih yang tak terhingga atas waktu, bimbingan, arahan, dan motivasi yang telah diberikan dengan penuh kesabaran selama proses penelitian hingga selesainya artikel ilmiah ini.
3. Ibu Prof. Dr. Kusriani, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta, atas kesempatan dan fasilitas yang telah diberikan selama menempuh studi.

4. Segenap dosen pengajar pada Program Studi SI Sistem Informasi Universitas Amikom Yogyakarta, atas ilmu, pendidikan, dan pengetahuan yang telah dibagikan kepada penulis selama menempuh perkuliahan.
5. Segenap staf dan pegawai Fakultas Ilmu Komputer Universitas Amikom Yogyakarta yang telah banyak membantu dalam urusan akademik ini.
6. Semua pihak yang tidak dapat di penulis sebutkan satu per satu, yang telah senantiasa mendukung baik secara langsung maupun tidak langsung. Semoga Allah SWT membalas segala kebaikan yang telah diberikan.

Yogyakarta, 13 January 2020

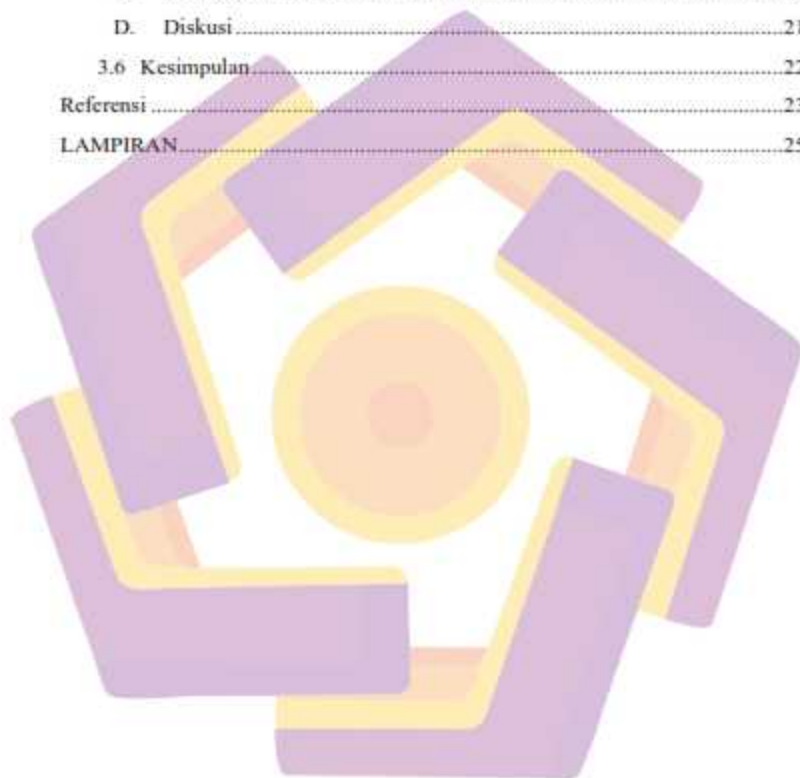
Penulis



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KONTRIBUSI KARYA.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
INTISARI	x
ABSTRACT.....	xi
BAB I IDENTITAS PUBLIKASI	1
BAB II PROSES SUBMIT.....	2
2.1 Lembar Review.....	2
2.2.1 Review Round 1	2
2.2.2 Review Round 2	3
2.2 Lembar Persetujuan (LoA).....	4
2.3 Sertifikat.....	6
2.4 Hasil Diskusi Ujian (Opsional).....	6
BAB III ISI KARYA ILMIAH.....	7
3.1 Intisari	7
3.2 Pendahuluan.....	7
3.3 Metode	12
A. Desain Penelitian	12
B. Dataset	13
C. Prapemrosesan Data.....	13
3.4 Analisis Kebutuhan dan Dampak Penerapan SMOTE	14
A. Algoritma pembelajaran Mesin	15

B. Metrik Evaluasi.....	16
3.5 Hasil dan Pembahasan	17
A. Matriks Konfusi.....	18
B. Evaluasi Kinerja.....	19
C. Grafik Kurva ROC.....	20
D. Diskusi.....	21
3.6 Kesimpulan.....	22
Referensi	23
LAMPIRAN.....	25



INTISARI

Untuk meningkatkan diagnosis diabetes yang lebih dini dan akurat, penelitian ini mengevaluasi tujuh model machine learning pada dataset Pima Indians Diabetes. Pada tahap awal, penelitian ini menangani masalah ketidakseimbangan kelas yang signifikan serta skala fitur data dengan menggunakan metode Synthetic Minority Over-sampling Technique (SMOTE) dan standarisasi z-score. Hasil analisis menunjukkan bahwa algoritma K-Nearest Neighbors (KNN) merupakan model yang paling efektif, dengan performa yang melampaui seluruh klasifikator lainnya. KNN memperoleh nilai recall yang sangat tinggi yaitu sebesar 94,06% serta nilai F1-score terbaik sebesar 0,8407. Hasil ini memiliki implikasi penting dalam praktik klinis karena nilai recall yang tinggi diperlukan untuk mengurangi jumlah kesalahan negatif palsu (false negatives) yang berbahaya dalam proses diagnosis. Meskipun model lain seperti SVM dan Random Forest juga menghasilkan performa yang baik, penelitian ini menegaskan bahwa pemilihan algoritma dengan sensitivitas tinggi seperti KNN, yang didukung oleh proses pra-pemrosesan data yang ketat, sangat penting dalam pengembangan sistem pendukung keputusan klinis yang andal.

Kata kunci—klasifikasi diabetes, pembelajaran mesin, SMOTE, dataset tidak seimbang.

ABSTRACT

To improve early and accurate diabetes diagnosis, we evaluated seven machine learning models on the Pima Indians Diabetes dataset. We first addressed the data's significant class imbalance and feature scaling issues using the Synthetic Minority Over-sampling Technique (SMOTE) and z-score standardization. Our analysis revealed that K-Nearest Neighbors (KNN) was the most effective model, outperforming all other classifiers. KNN had a great recall rate of 94.06% and the best F1-score of 0.8407. This result is important for clinical practice because high recall is necessary to reduce the number of dangerous false negatives in a diagnostic setting. Although other models such as SVM and Random Forest also yielded commendable results, we assert that the selection of a high-sensitivity algorithm like KNN, in conjunction with rigorous data preprocessing, is essential for the creation of reliable clinical decision support systems.

Keywords—diabetes classification, machine learning, SMOTE, imbalanced dataset