



DETEKSI PENYAKIT DIABETES MENGGUNAKAN METODE *K-NEAREST NEIGHBORS* (K-NN) DAN SELEKSI FITUR *CHI-SQUARE*

SKRIPSI

**Diajukan Sebagai Syarat untuk Menyelesaikan
Pendidikan Program Sata-1 Pada
Program Studi Teknik Informatika**

Oleh:

**Rafif Radhi Rachman
2021.11.0124**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN SAINS
UNIVERSITAS INDO GLOBAL MANDIRI
2026**

DETEKSI PENYAKIT DIABETES MENGGUNAKAN METODE *K-NEAREST NEIGHBORS* (K-NN) DAN SELEKSI FITUR *CHI-SQUARE*

SKRIPSI



**Diajukan Sebagai Syarat untuk Menyelesaikan
Pendidikan Program Sata-1 Pada
Program Studi Teknik Informatika**

Oleh:

Rafif Radhi Rachman

2021110124

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN SAINS
UNIVERSITAS INDO GLOBAL MANDIRI**

2026

LEMBAR PENGESAHAN SKRIPSI

LEMBAR PENGESAHAN SKRIPSI

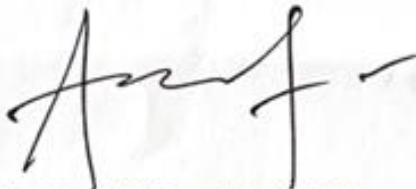
**Deteksi Penyakit Diabetes Menggunakan Metode K-Nearest
Neighbors (K-NN) Dan Seleksi Fitur Chi-Square**

Oleh

Rafif Radhi Rachman
NPM : 2021.11.0124

Palembang 27 Juli 2026

Pembimbing I



Dr. Lastri Widya Astuti, M.Kom
NIK. 2003.01.0063

Pembimbing II



Dr. Muhammad Haviz Irfani, S.Si., M.T.
NIK. 2021.03.0291

Mengetahui,

Dekan Fakultas Ilmu Komputer dan Sains

FAKULTAS ILMU KOMPUTER & SAINS



Dr. Herri Setiawan, M.Kom.
NIK. 2003.01.0060

LEMBAR PERSETUJUAN DEWAN PENGUJI

LEMBAR PERSETUJUAN DEWAN PENGUJI

Pada hari Jum'at tanggal 13 Juli 2026 telah dilaksanakan ujian sidang skripsi :

Nama : Rafif Radhi Rachman
NPM : 2021.11.0124
Judul : Deteksi Penyakit Diabetes Menggunakan Metode K-Nearest
Neighbors (K-NN) Dan Seleksi Fitur Chi-Square

Oleh Prodi Teknik Informatika Fakultas Ilmu Komputer dan Sains Universitas
Indo Global Mandiri Palembang

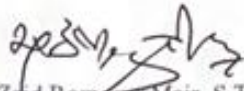
Palembang 27 Juli 2026

Penguji 1,



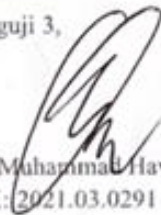
Rudi Heriansyah, S.T., M.Eng., Ph.D
NIK: 2022.01.0315

Penguji 2,



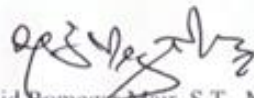
Zaid Romegat Mair, S.T., M.Cs
NIK: 2021.01.0307

Penguji 3,



Dr. Muhammad Haviz Irfani, S.Si., K.T.I
NIK: 2021.03.0291

Menyetujui,
Ka. Prodi Teknik Informatika



Zaid Romegat Mair, S.T., M.Cs
NIK: 2021.01.0307

SURAT KETERANGAN REVISI SKRIPSI



SURAT KETERANGAN REVISI SKRIPSI
PROGRAM STUDI TEKNIK INFORMATIKA (SI)
FASILKOM DAN SAINS UNIVERSITAS INDO GLOBAL MANDIRI

Kami yang bertanda tangan dibawah ini, menerangkan bahwa :

Nama : Rafif Radhi Rachman
NPM : 2021.11.0124
Judul : Deteksi Penyakit Diabetes Menggunakan Metode K-Nearest
Neighbors (K-NN) Dan Seleksi Fitur Chi-Square

Mahasiswa yang namanya tercantum diatas, telah selesai merevisi penulisan SKRIPSI

Palembang, 27 Juli 2026.

Penguji 1,

Rudi Heriansyah, S.T., M.Eng., Ph.D
NIK: 2022.01.0315

Penguji 2,

Zaid Romegar Mair, S.T., M.Cs
NIK: 2021.01.0307

Penguji 3,

Dr. Muhammad Hayiz Irfani, S.Si., K.T.I
NIK: 2021.03.0291

Menyetujui,
Ka. Prodi Teknik Informatika

Zaid Romegar Mair, S.T., M.Cs
NIK: 2021.01.0307

MOTTO DAN PERSEMBAHAN

❖ Motto

Di balik setiap ujian yang berat, selalu ada pelajaran yang mendewasakan. Tidak ada kesuksesan tanpa perjuangan, dan tidak ada kemudahan tanpa melewati kesulitan terlebih dahulu.

❖ Persembahan

Dengan penuh rasa syukur dan ketulusan, skripsi ini saya persembahkan kepada:

1. Allah SWT yang telah memberikan kemudahan, kebaikan, serta kesuksesan dalam setiap langkah hidupku, dan Rasulullah Muhammad SAW sebagai teladan hidup yang sempurna.
2. Kedua Orang tua tercinta yang selalu memberikan doa, dukungan, dan semangat tanpa henti. Berjuang tanpa lelah memperjuangkan pendidikan penulis sampai berada di titik ini tanpa kenal lelah. Tak terkira ucapan terima kasih atas kesempatan yang kedua orang tua berikan kepada penulis hingga bisa berada pada tahap ini. Semoga usaha dan upaya yang telah dilakukan selama ini dapat menjadi berkah untuk kedepannya.
3. Dosen Pembimbing, Ibu Lastri Widya Astuti, M.Kom dan Bapak Muhammad Haviz Irfani, S.Si., M.T.I. Terima kasih atas kesabaran, bimbingan, dan waktu yang telah diberikan selama proses pengerjaan skripsi ini. Tanpa arahan dari kalian, perjalanan ini tidak akan terselesaikan dengan baik.
4. Bapak/Ibu Dosen Teknik Informatika dan seluruh Civitas Akademika Universitas Indo Global Mandiri, yang telah menjadi bagian penting dalam perjalanan akademisku.
5. Untuk teman dan sahabat tercinta , terima kasih sudah kebersamaan penulis hingga sampai di titik ini. Terima kasih untuk semua dukungan dan pertolongan yang telah kalian berikan, menemani di setiap proses yang penulis lewati. Semoga

kebahagiaan akan selalu menyertai kalian dan semoga persahabatan ini tidak lekang di makan usia.

6. Teman-teman seperjuangan angkatan 2021, yang telah menjadi rekan dalam belajar, berbagi tawa, serta menghadapi berbagai tantangan selama masa studi. Terima kasih telah menjadi bagian dari perjalanan ini.
7. Dan yang terakhir, untuk diriku sendiri, Rafif Radhi Rachman. Terima kasih telah berjuang sekeras ini, menahan seikhlas ini dan bertahan sejauh ini. Meskipun jalan didepan masih panjang dan akan lebih berat, mari rayakan saat ini untuk mengenang usaha yang telah di perjuangkan.

Semoga setiap langkah yang telah diambil dan setiap usaha yang telah dilakukan menjadi awal dari perjalanan yang lebih besar di masa depan.

DETEKSI PENYAKIT DIABETES MENGGUNAKAN METODE K-NEAREST NEIGHBORS (K-NN) DAN SELEKSI FITUR CHI-SQUARE

ABSTRAK

Diabetes Mellitus (DM) merupakan penyakit metabolik kronis yang prevalensinya terus meningkat, khususnya pada kelompok lansia. Gejala yang tidak spesifik pada lansia sering menyebabkan keterlambatan *diagnosis* sehingga diperlukan sistem deteksi dini berbasis *machine learning*. Penelitian ini menerapkan algoritma *K-Nearest Neighbors* (K-NN) dikombinasikan dengan seleksi fitur *Chi-Square* untuk mendeteksi Diabetes Mellitus Tipe 2 pada pasien lansia di RS PUSRI Palembang menggunakan 233 data rekam medis. Seleksi fitur *Chi-Square* menghasilkan lima fitur paling relevan yaitu Kadar Glukosa (111,32), Kadar Hemoglobin A1c (101,74), Kadar Kolesterol (3,87), Riwayat Keluarga (3,73), dan Umur (2,82). Hasil evaluasi menunjukkan model K-NN tanpa seleksi fitur mencapai akurasi 62,5%, sedangkan dengan seleksi fitur *Chi-Square* meningkat signifikan menjadi akurasi 83,3%, presisi 82,4%, *recall* 93,3%, dan *F1-Score* 87,5%. Hal ini membuktikan bahwa seleksi fitur *Chi-Square* mampu meningkatkan performa klasifikasi K-NN dalam deteksi dini diabetes pada lansia.

Kata Kunci: Diabetes Mellitus, *K-Nearest Neighbors*, *Chi-Square*, Seleksi Fitur, Lansia

DIABETES DISEASE DETECTION USING THE K-NEAREST NEIGHBORS (K-NN) METHOD AND CHI-SQUARE FEATURE SELECTION

ABSTRACT

Diabetes Mellitus (DM) is a chronic metabolic disease with a steadily increasing prevalence, particularly among the elderly population. Non-specific symptoms in older patients frequently lead to delayed diagnosis, necessitating a machine learning-based early detection system. This study applies the K-Nearest Neighbors (K-NN) algorithm combined with Chi-Square feature selection to detect Type 2 Diabetes Mellitus in elderly patients at PUSRI Hospital Palembang using 233 medical records. Chi-Square feature selection identified five most relevant features: Blood Glucose Level (111.32), HbA1c Level (101.74), Cholesterol Level (3.87), Family History (3.73), and Age (2.82). Evaluation results show that the K-NN model without feature selection achieved an accuracy of 62.5%, while the model with Chi-Square feature selection improved significantly to an accuracy of 83.3%, precision of 82.4%, recall of 93.3%, and F1-Score of 87.5%. These results confirm that Chi-Square feature selection effectively enhances K-NN classification performance for early diabetes detection in elderly patients.

Keywords: Diabetes Mellitus, K-Nearest Neighbors, Chi-Square, Feature Selection, Elderly

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan hidayah-Nya sehingga penelitian ini dapat diselesaikan dengan baik dan tepat pada waktunya. Shalawat serta salam semoga selalu tercurah kepada Nabi Muhammad SAW, keluarga, para sahabat, dan kita semua sebagai pengikutnya hingga akhir zaman.

Skripsi dengan judul “Deteksi Penyakit Diabetes Menggunakan Metode *K-Nearest Neighbors* (K-NN) Dan Seleksi Fitur *Chi-Square*” disusun untuk memenuhi salah satu syarat dalam menyelesaikan pendidikan jenjang Sarjana (S1) pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer dan Sains, Universitas Indo Global Mandiri (UIGM) Palembang. Penulis menyadari bahwa penyelesaian penelitian ini tidak akan terlaksana tanpa dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Marzuki Alie, SE., MM, selaku Rektor Universitas Indo Global Mandiri (UIGM).
2. Bapak Plt. Dr. Herri Setiawan, M.Kom, selaku Dekan Fakultas Ilmu Komputer dan Sains Universitas Indo Global Mandiri (UIGM).
3. Bapak Zaid Romegar Mair, S.T., M.Cs, selaku Ketua Program Studi Teknik Informatika Universitas Indo Global Mandiri (UIGM).
4. Ibu Lastri Widya Astuti, M.Kom selaku Dosen Pembimbing I, yang telah memberikan bimbingan, arahan, dan masukan berharga dalam proses penyelesaian skripsi ini.
5. Bapak Muhammad Haviz Irfani, S.Si., M.T.I, selaku Dosen Pembimbing II, yang telah memberikan bimbingan, arahan, dan masukan berharga dalam proses penyelesaian skripsi ini.
6. Seluruh Bapak/Ibu Dosen Fakultas Ilmu Komputer dan Sains, yang telah memberikan ilmu dan pengalaman yang sangat bermanfaat selama masa studi.

7. Orang tua tercinta yang selalu memberikan doa, dukungan, dan semangat tanpa henti.
8. Teman-teman seperjuangan, yang telah membantu dan mendukung penulis selama menyusun skripsi ini, baik secara langsung maupun tidak langsung.

Penulis menyadari bahwa penelitian ini masih jauh dari sempurna. Oleh karena itu, penulis sangat terbuka terhadap kritik dan saran yang membangun sebagai bahan perbaikan di masa yang akan datang. Semoga skripsi ini dapat memberikan manfaat, khususnya dalam bidang ilmu teknologi informasi dan secara umum bagi siapa saja yang membutuhkan.

Palembang, 27 April 2026

Penulis,

Rafif Radhi Rachman

2021110124

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	iii
LEMBAR PERSETUJUAN DEWAN PENGUJI.....	iv
SURAT KETERANGAN REVISI SKRIPSI	v
ABSTRAK	viii
<i>ABSTRACT</i>	ix
KATA PENGANTAR.....	x
DAFTAR ISI.....	xiii
DAFTAR GAMBAR	xvii
DAFTAR TABEL.....	xviii
DAFTAR LAMPIRAN	xix
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan dan Manfaat Penelitian	4
1.4.1 Tujuan Penelitian.....	4
1.4.2 Manfaat Penelitian	5
1.4.3 Metodologi Penelitian	5
1.5 Sistematika Penulisan	6
BAB II.....	7
LANDASAN TEORI.....	7

2.1	Rumah Sakit	7
2.1.1	Rekam Medis	8
2.2	Penyakit Diabetes Melitus	8
2.2.1	Diabetes Pada Lansia	9
2.2.2	Hasil Laboratorium Diabetes	10
2.3	Machine Learning	12
2.3.1	Algoritma Klasifikasi <i>K-Nearest Neighbors</i> (K-NN)	13
2.3.2	Metode Seleksi Fitur Chi-Square	15
2.4	Pemrograman	17
2.4.1	Pemrograman Python	17
2.5	Evaluasi Matrix	19
2.5.1	<i>Confusion Matrix</i>	19
2.6	Penelitian Terdahulu	22
BAB III		26
METODOLOGI PENELITIAN		26
3.1	Tahapan Penelitian	26
3.2	Observasi	27
3.3	Studi Literatur	28
3.4	Pengumpulan Data	29
3.5	Pra-Pemrosesan Data	32
3.5.1	Penghapusan Kolom yang Tidak Digunakan	32
3.5.2	Penghapusan Data Duplikat	34
3.5.3	Diskretisasi Data	34
3.5.4	Split Data	38
3.6	Implementasi Algoritma	38

3.6.1 Penerapan Algoritma Chi-Square	39
3.6.2 Penerapan Metode K-NN.....	49
3.7 Pengujian.....	67
3.7.1 Pengujian K-NN Tanpa Seleksi Fitur.....	67
3.7.2 Pengujian K-NN Dengan Seleksi Fitur	68
3.8 Evaluasi.....	70
3.8.1 Evaluasi K-NN Tanpa Seleksi Fitur	70
3.8.2 Evaluasi K-NN Dengan Seleksi Fitur	71
3.9 Kesimpulan	71
BAB IV	73
HASIL DAN PEMBAHASAN	73
4.1 Pengumpulan Data	73
4.2 Pra-Pemrosesn Data	75
4.2.1 Penghapusan Kolom Yang Tidak Digunakan.....	75
4.2.2 Penghapusan Data Duplikat.....	79
4.2.3 Diskretisasi Data	80
4.2.4 Split Data.....	84
4.3 Implementasi Algoritma.....	85
4.3.1 Implementasi Algoritma Chi-Square	85
4.3.2 Implementasi Algoritma K-NN.....	96
4.4 Pengujian.....	97
4.5 Hasil	105
4.6 Pembahasan.....	110
BAB V.....	112
KESIMPULAN DAN SARAN.....	112

5.1 Kesimpulan	112
5.2 Saran.....	112
DAFTAR PUSTAKA	114
LAMPIRAN.....	123

DAFTAR GAMBAR

Gambar 3.1 Tahapan penelitian	27
Gambar 3.2 Observasi Di Rumah Sakit PUSRI Palembang	29
Gambar 3.3 Tahapan Pra-Pemrosesan Data.....	32
Gambar 3.4 Hasil Penghapusan Data Duplikat	34
Gambar 3.5 Tahapan Implementasi Algoritma	39
Gambar 3.6 Tahapan Penerapan Metode K-NN	49
Gambar 3.7 <i>Confusion Matrix</i> Tanpa Seleksi Fitur	68
Gambar 3.8 <i>Confusion Matrix</i> Dengan Seleksi Fitur	69
Gambar 4.1 Dokumentasi Pengumpulan Data	73
Gambar 4.2 Grafik Hasil Percobaan K3	102
Gambar 4.3 Grafik Hasil Percobaan K5	103
Gambar 4.4 Grafik Hasil Percobaan K7	104
Gambar 4.5 <i>Confusion Matrix</i> Best Model K3	107
Gambar 4.6 <i>Confusion Matrix</i> Best Model K5	107
Gambar 4.7 <i>Confusion Matrix</i> Best Model K7	108

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	20
Tabel 2.2 Penelitian Terdahulu	23
Tabel 3.1 Data Pasien Penyakit Diabetes RS. Pusri Palembang	31
Tabel 3.2 Hasil Penghapusan Kolom Yang Tidak Digunakan	33
Tabel 3.3 Hasil Diskretisasi Data	37
Tabel 3.4 Hasil Kontingensi	41
Tabel 3.5 Hasil Perhitungan <i>Expected Frequency</i>	44
Tabel 3.6 Hasil Perhitungan Nilai <i>Chi-Square</i>	46
Tabel 3.7 Hasil Perankingan Nilai <i>Chi-Square</i>	48
Tabel 3.8 Contoh Hasil Tetangga Terdekat Pada $K=4$	51
Tabel 3.9 Perbandingan Nilai K tanpa Seleksi Fitur	51
Tabel 3.10 Data latih K -NN Tanpa Seleksi Fitur	53
Tabel 3.11 Data uji K -NN Tanpa Seleksi Fitur	54
Tabel 3.12 Hasil Perhitungan K -NN Tanpa Seleksi Fitur	56
Tabel 3.13 Hasil Pemeringkatan Tetangga Terdekat Tanpa Seleksi Fitur	58
Tabel 3.14 Perbandingan Nilai K dengan Seleksi Fitur	60
Tabel 3.15 Perbandingan Nilai K Dengan Seleksi Fitur	61
Tabel 3.16 Hasil Pemeringkatan Tetangga Terdekat Dengan Seleksi Fitur	63
Tabel 3.16 Hasil Pemeringkatan Tetangga Terdekat Dengan Seleksi Fitur	63
Tabel 4.1 Data Pasien Penyakit Diabetes RS. Pusri Palembang	74
Tabel 4.2 Hasil Penghapusan Kolom Yang Tidak Digunakan	77
Tabel 4.3 Hasil Diskretisasi Data	82
Tabel 4.4 Hasil Kontigensi	86

Tabel 4.5 Hasil Perhitungan Expected Frequency	89
Tabel 4.6 Hasil Perhitungan Nilai Chi-Square	91
Tabel 4.7 Hasil Perankingan Nilai Chi-Square	93
Tabel 4.8 Hasil Proses TOP Fitur	95
Tabel 4.9 Hasil Proses Pengujian K3	99
Tabel 4.10 Hasil Proses Pengujian K5	100
Tabel 4.11 Hasil Proses Pengujian K7	101
Tabel 4.12 Hasil Proses Pemilihan 3 Model Terbaik (K3, K5, K7)	105

DAFTAR LAMPIRAN

Lampiran 1. 1 Daftar Riwayat Hidup	122
Lampiran 1. 2 Lampiran Kartu Bimbingan/Konsultasi.....	123
Lampiran 1. 3 Surat Pernyataan Tidak Plagiat	124