

SKRIPSI

**OPTIMASI ALGORITMA K-NEAREST NEIGHBORS PADA
PREDIKSI PENYAKIT DIABETES**

***OPTIMIZATION OF THE K-NEAREST NEIGHBORS
ALGORITHM IN DIABETES PREDICTION***

Untuk Memenuhi Sebagian Persyaratan
Memperoleh Gelar Sarjana Komputer



**DISUSUN OLEH:
SITTI ARFIAH
D0221010**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS SULAWESI BARAT
MAJENE**

2025

SKRIPSI

**OPTIMASI ALGORITMA K-NEAREST NEIGHBORS PADA
PREDIKSI PENYAKIT DIABETES**

*Optimization of the K-Nearest Neighbors Algorithm in Diabetes
Prediction*

Untuk Memenuhi Sebagian Persyaratan
Memperoleh Gelar Sarjana Komputer



**DISUSUN OLEH:
SITTI ARFIAH
D0221010**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS SULAWESI BARAT
MAJENE
2025**

LEMBAR PERSETUJUAN

SKRIPSI

OPTIMASI ALGORITMA K-NEAREST NEIGHBORS PADA PREDIKSI PENYAKIT DIABETES

Telah dipersiapkan dan disusun oleh

SITTI ARFIAH

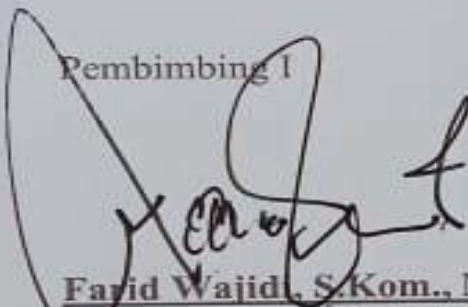
NIM: D0221010

Telah dipertahankan di depan Tim Penguji

Pada tanggal 06 November 2025

Susunan Tim Penguji

Pembimbing I



Farid Wajid, S.Kom., M.T

NIP. 198904182019031018

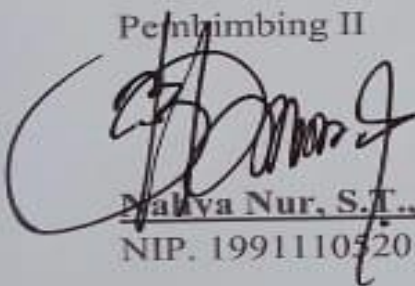
Penguji I



Dr. Eng. Sulfayanti, S.Si., M.T

NIP. 198903172020122011

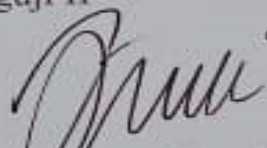
Pembimbing II



Mahya Nur, S.T., M.Kom

NIP. 199111052019032024

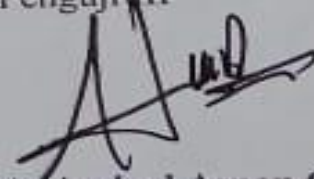
Penguji II



Nurhikma Arifin, S.Kom., M.T

NIP. 199304252022032011

Penguji III



Amirul Asnan Cirua, S.T., M.Kom

NIP. 199804022024061001

LEMBAR PENGESAHAN

OPTIMASI ALGORITMA K-NEAREST NEIGHBORS PADA
PREDIKSI PENYAKIT DIABETES

SKRIPSI

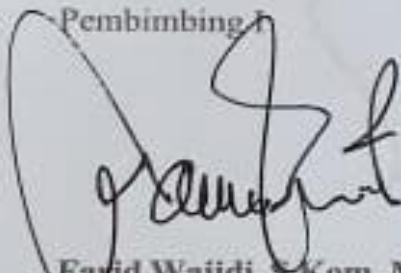
Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun oleh:

SITTI ARFIAH
NIM: D0221010

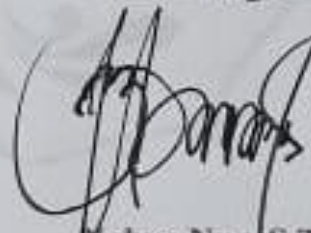
Telah disetujui oleh dosen pembimbing Skripsi
Pada Tanggal 06 November 2025

Pembimbing I



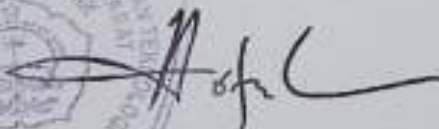
Fayid Wajidi, S.Kom., M.T.
NIP: 198904182019031018

Pembimbing II



Nahya Nur, S.T., M.Kom
NIP: 199111052019032024

Dekan Fakultas Teknik,
Universitas Sulawesi Barat



Prof. Dr. Ir. Hafsah Nirwana, M.T
NIP: 196404051990032002

Ketua Program Studi
Informatika,



Muh Rafli Rasyid, S. Kom., M.T
NIP: 19880818202231006

PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar referensi.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur unsur plagiaris, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70)

Majene, 01 Juli 2025



METERAI
TEMPEL
10000
BFDCFANX101987199

Sitti Arfiah
NIM: D0221010

ABSTRAK

Diabetes melitus merupakan penyakit kronis yang ditandai dengan tingginya kadar gula darah akibat gangguan pada fungsi insulin. Jika tidak ditangani dengan tepat, penyakit ini dapat menyebabkan komplikasi serius seperti kerusakan pada jantung, ginjal, dan saraf. Oleh karena itu, deteksi dini sangat penting. Penelitian ini bertujuan meningkatkan akurasi klasifikasi diabetes menggunakan algoritma *K-Nearest Neighbors* (KNN). Pada model awal, KNN diterapkan tanpa optimasi dan penyeimbangan data, menghasilkan *accuracy* sebesar 83%. Namun, performa terhadap kelas positif masih rendah, menunjukkan bias model terhadap kelas mayoritas. Untuk mengatasi hal ini, dilakukan beberapa tahapan, yaitu preprocessing data, optimasi *hyperparameter* menggunakan *GridSearchCV*, serta penyeimbangan data dengan metode SMOTE. Setelah penerapan ketiga tahapan tersebut, *accuracy* meningkat menjadi 94%, dengan peningkatan signifikan pada performa kelas positif (*precision* 90%, *recall* 98%, dan *F1-score* 94%). Sementara itu, kelas negatif juga menunjukkan *accuracy* sebesar 94% , *precision* 98%, *recall* 89%, *F1-score* sebesar 93%. Hasil penelitian ini menunjukkan bahwa kombinasi *preprocessing*, optimasi model, dan penyeimbangan kelas dapat meningkatkan kinerja algoritma KNN secara signifikan dalam mendeteksi diabetes secara lebih adil dan akurat. Pendekatan ini dapat diterapkan dalam pengembangan sistem pendukung keputusan medis, khususnya dalam proses diagnosis dini penyakit diabetes

Kata Kunci: Optimasi K-Nearest Neighbors; GridSearchCV; Synthetic Minority Over-sampling Technique; Prediksi Diabetes

ABSTRACT

Diabetes mellitus is a chronic disease characterized by elevated blood glucose levels due to impaired insulin function. If not properly managed, this condition can lead to serious complications such as damage to the heart, kidneys, and nerves. Therefore, early detection is crucial. This study aims to improve the accuracy of diabetes classification using the K-Nearest Neighbors (KNN) algorithm. In the initial model, KNN was applied without any optimization or data balancing, resulting in an accuracy of 83%. However, the performance on the positive class remained low, indicating the model's bias toward the majority class. To address this issue, several steps were implemented: data preprocessing, hyperparameter optimization using GridSearchCV, and class balancing using the SMOTE method. After applying these steps, the model's accuracy increased to 94%, with a significant improvement in the performance of the positive class (precision 90%, recall 98%, and F1-score 94%). Meanwhile, the negative class also showed strong results with an accuracy of 94%, precision 98%, recall 89%, and F1-score 93%. These findings demonstrate that the combination of preprocessing, model optimization, and class balancing can significantly enhance the performance of the KNN algorithm in detecting diabetes in a more fair and accurate manner. This approach can be applied in the development of medical decision support systems, particularly in the early diagnosis of diabetes.

Keywords: *Optimization K-Nearest Neighbors; GridSearchCV; Synthetic Minority Over-sampling Technique; Diabetes Prediction*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Diabetes melitus merupakan masalah kesehatan global yang terus meningkat. Menurut *International Diabetes Federation (IDF, 2021)* , terdapat sekitar 537 juta penderita diabetes di dunia dan diperkirakan meningkat menjadi 783 juta pada tahun 2045. Indonesia sendiri menempati urutan kelima terbesar dengan sekitar 19,47 juta penderita pada tahun 2021. Kondisi ini menunjukkan bahwa sistem kesehatan perlu bertransformasi untuk menangani penyakit yang semakin besar, terutama dalam upaya pencegahan dan penanganan. Data *Survei Kesehatan Indonesia (SKI) 2023* dan *Riskesmas 2018* juga menampilkan tren peningkatan prevalensi diabetes, khususnya Diabetes Melitus Tipe 2 (DMT2) yang mencapai 10,6% pada kelompok usia 20–79 tahun. Peningkatan kasus ini terutama terjadi di wilayah perkotaan akibat gaya hidup tidak sehat, obesitas, dan faktor risiko lainnya (Kementerian Kesehatan Republik Indonesia, 2025).

Menurut definisi yang diberikan oleh *World Health Organization (WHO)*, kesehatan tidak hanya diartikan sebagai ketiadaan penyakit atau kondisi lemah fisik semata, melainkan mencakup suatu keadaan yang utuh dan menyeluruh, mencakup kesejahteraan fisik, kondisi mental yang stabil, serta kehidupan sosial yang harmonis. Dengan kata lain, seseorang dapat dikatakan benar-benar sehat apabila ia berada dalam kondisi seimbang dan optimal pada aspek ketiga tersebut, bukan hanya karena tidak mengalami gangguan kesehatan secara fisik (Najah, 2022).

Seiring berjalannya waktu, struktur masyarakat mengalami perubahan yang signifikan. Perubahan ini dipicu oleh pergeseran dari kehidupan agraris menuju era industri. Dampaknya terlihat pada pola konsumsi makanan dan tingkat aktivitas fisik masyarakat. Misalnya kebiasaan makan telah berubah, dengan semakin banyaknya orang yang memilih makanan cepat saji karena kepraktisannya dan kemudahannya aksesnya. Gaya hidup yang berkembang saat ini telah menyebabkan penurunan signifikan dalam aktivitas fisik, terutama di kalangan pekerja kantoran yang cenderung menghabiskan sebagian

besar waktu mereka di dalam ruangan dengan sedikit gerakan. Faktor – faktor ini berperan penting dalam meningkatkan risiko penyakit tidak menular serta penyakit degeneratif. Salah satu penyakit tidak menular yang sering dikaitkan dengan kebiasaan makan yang buruk dan kurangnya aktivitas fisik adalah diabetes melitus (Astutisari et al., 2022).

Diabetes melitus terjadi akibat gangguan dalam proses metabolisme tubuh, yang menyebabkan peningkatan kadar glukosa dalam darah, padahal glukosa adalah sumber energi utama bagi sel-sel tubuh. Namun, apabila kadar glukosa tidak terjaga dengan baik, kondisi ini dapat menyebabkan komplikasi serius seperti penyakit kardiovaskuler, stroke, obesitas, serta gangguan pada mata, ginjal, dan sistem saraf (Argina, 2020). Salah satu bentuk diabetes yang paling sering ditemui adalah diabetes melitus (Hartono, 2024). Diabetes melitus sendiri merupakan gangguan metabolik yang ditandai dengan kadar gula darah yang tinggi (hiperglikemia), yang disebabkan oleh gangguan dalam produksi insulin, kemampuan insulin bekerja, atau keduanya (Fandinata & Darmawan, 2020). Salah satu algoritma yang banyak digunakan dalam mengoptimasi prediksi penyakit diabetes adalah *K-Nearest Neighbors* (KNN). Pemilihan algoritma KNN dalam penelitian ini didasarkan pada kemampuannya mengklasifikasikan data berdasarkan kedekatan nilai antar fitur, yang sangat relevan dengan data medis seperti kadar glukosa, tekanan darah, dan BMI. Seperti yang diketahui, diabetes melitus merupakan gangguan metabolik yang cukup kompleks, ditandai dengan tingginya kadar glukosa dalam darah dan berpotensi menimbulkan komplikasi serius jika tidak segera ditangani (Argina, 2020). Sifatnya yang non-parametrik membuat KNN fleksibel dalam menangani data dengan banyak variabel seperti glukosa, tekanan darah, dan faktor-faktor metabolik lainnya (Fandinata & Darmawan, 2020). Namun, performa KNN sangat dipengaruhi oleh parameter yang digunakan, seperti jumlah tetangga (k), bobot, dan metrik jarak. Jika parameter ini tidak ditentukan secara optimal, model dapat mengalami penurunan *accuracy* dan sensitivitas, terutama terhadap kelas minoritas. Oleh karena itu, dilakukan optimasi menggunakan *GridSearchCV* untuk memperoleh

kombinasi parameter terbaik yang mampu meningkatkan kinerja model, khususnya dalam mendeteksi kasus positif diabetes.

Selain itu, salah satu tantangan utama dalam prediksi penyakit seperti diabetes adalah ketidakseimbangan kelas pada dataset, di mana jumlah data penderita dan non-penderita tidak seimbang. Untuk mengatasi hal tersebut, digunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE), yang terbukti efektif dalam memperbaiki distribusi data dan meningkatkan kinerja model klasifikasi. Ketidakseimbangan kelas dapat menyebabkan model lebih cenderung memprediksi kelas mayoritas, sehingga mengurangi sensitivitas terhadap kelas minoritas yang justru sangat penting dalam konteks prediksi penyakit. Oleh karena itu, KNN dan SMOTE dipilih dalam penelitian ini sebagai dasar pengembangan model prediksi, dengan harapan mampu memberikan hasil klasifikasi yang akurat dan mudah dioptimalkan, serta penanganan ketidakseimbangan pada dataset.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana hasil penerapan Optimasi Algoritma *K-Nearest Neighbor* (KNN) pada prediksi penyakit Diabetes ?
2. Bagaimana hasil penerapan *Synthetic Minority Over-sampling Technique* (SMOTE) dan *GridSearchCV* pada prediksi penyakit Diabetes ?

1.3 Batasan Masalah

Agar penelitian ini lebih terfokus dan terarah, maka batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah *Pima Indian Diabetes Dataset* yang bersumber dari Kaggle.
2. Penelitian ini hanya berfokus pada Optimasi Algoritma *K-Nearest Neighbor* (KNN) pada prediksi penyakit Diabetes, tanpa membandingkan dengan metode lain.

3. Optimasi model Algoritma *K-Nearest Neighbor* (KNN) pada prediksi penyakit Diabetes dilakukan menggunakan dua pendekatan spesifik yaitu *Synthetic Minority Over-sampling Technique* (SMOTE) untuk menangani ketidakseimbangan kelas, dan *GridSearchCV* untuk menemukan kombinasi *hyperparameter* terbaik.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

1. Untuk mengetahui hasil penerapan Optimasi Algoritma *K-Nearest Neighbor* (KNN) pada prediksi penyakit Diabetes.
2. Untuk mengetahui hasil penerapan *Synthetic Minority Over-sampling Technique* (SMOTE) dan *GridSearchCV* untuk menangani ketidakseimbangan data pada prediksi penyakit diabetes.

1.5 Manfaat Penelitian

Penelitian ini bermanfaat untuk membantu peningkatan kinerja yang besar dalam mendeteksi kasus positif (penderita diabetes), yang sebelumnya cenderung terabaikan pada dataset yang tidak seimbang. Hal ini sangat penting karena deteksi dini dan akurat memungkinkan penderita memperoleh penanganan medis yang tepat. Serta Hasil penelitian ini dapat berkontribusi dalam membangun sistem deteksi dini yang lebih efektif dan akurat untuk penyakit diabetes, sehingga dapat mengurangi dampak global dari penyakit ini.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil evaluasi, penerapan optimasi algoritma *K-Nearest Neighbors* (KNN) dengan *GridSearchCV* dan penyeimbangan data menggunakan SMOTE memberikan peningkatan kinerja yang signifikan. Pada model awal dengan data tidak seimbang, *accuracy* hanya mencapai 83%, dengan performa kelas positif yang masih rendah (*precision* 80%, *recall* 69%, *F1-score* 74%) serta ketimpangan performa pada kelas negatif (*precision* 84%, *recall* 91%, *F1-score* 88%), menunjukkan adanya bias terhadap kelas negatif. Setelah dilakukan penyeimbangan data menggunakan SMOTE dan optimasi parameter, *accuracy* meningkat menjadi 94%, dengan perbaikan drastis pada kelas positif (*precision* 90%, *recall* 98%, *F1-score* 94%) dan performa tinggi juga pada kelas negatif (*precision* 98%, *recall* 89%, *F1-score* 93%). Penelitian ini menunjukkan bahwa kombinasi SMOTE dan optimasi parameter mampu meningkatkan kemampuan model dalam mengenali kedua kelas secara seimbang, khususnya dalam mendeteksi kasus positif yang krusial dalam prediksi penyakit diabetes.

5.2 Saran

Berdasarkan hasil evaluasi yang telah dilakukan, disarankan agar penelitian selanjutnya mempertimbangkan penggunaan teknik normalisasi yang berbeda seperti *StandardScaler* atau *RobustScaler* untuk dibandingkan dengan *MinMaxScaler*, guna mengetahui pengaruhnya terhadap performa model, khususnya dalam menghadapi data dengan outlier. Selain itu, meskipun model KNN menunjukkan performa yang baik setelah dilakukan penyeimbangan dan optimasi, akan lebih baik jika dilakukan perbandingan dengan algoritma lain seperti Logistic Regression, Decision Tree, Random Forest, atau Support Vector Machine (SVM) sebagai pembanding (baseline model) untuk memastikan bahwa KNN merupakan pilihan yang paling optimal dalam kasus ini. Disarankan juga untuk menggunakan teknik validasi silang seperti K-Fold Cross Validation guna memperoleh hasil evaluasi yang lebih

stabil dan tidak bergantung pada satu kali pembagian data. Peneliti juga sebaiknya melakukan analisis lebih lanjut terhadap kesalahan klasifikasi, terutama false negative, karena dalam konteks prediksi penyakit seperti diabetes, kesalahan ini dapat berakibat fatal. Visualisasi hasil evaluasi seperti confusion matrix, ROC curve, dan distribusi data sebelum dan sesudah SMOTE juga dapat memperkuat pemahaman terhadap performa model.

DAFTAR PUSTAKA

- Anggrawan, A., & Mayadi, M. (2023). Application of KNN Machine Learning and Fuzzy C-Means to Diagnose Diabetes. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 22(2), 405–418. <https://doi.org/10.30812/matrik.v22i2.2777>
- Argina, A. M. (2020). Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes. *Indonesian Journal of Data and Science*, 1(2), 29–33. <https://doi.org/10.33096/ijodas.v1i2.11>
- Astutisari, I. D. A. E. C., AAA Yulianti Darmini, A. Y. D., & Ida Ayu Putri Wulandari, I. A. P. W. (2022). Hubungan Pola Makan Dan Aktivitas Fisik Dengan Kadar Gula Darah Pada Pasien Diabetes Melitus Tipe 2 Di Puskesmas Manggis I. *Jurnal Riset Kesehatan Nasional*, 6(2), 79–87. <https://doi.org/10.37294/jrkn.v6i2.350>
- Cahyaputra, H. R., & Rahmadewi, R. (2024). Klasifikasi Tingkat Kematangan Buah Paprika Menggunakan Metode K-Nearest Neighbor Berdasarkan Warna Rgb Melalui Aplikasi Matlab. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 9(1), 242–249. <https://doi.org/10.29100/jupi.v9i1.4440>
- Cholil, S. R., Handayani, T., Prathivi, R., & Ardianita, T. (2021). Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 6(2), 118–127. <https://doi.org/10.31294/ijcit.v6i2.10438>
- Clara, S., Laksmi Prianto, D., Al Habsi, R., Friscila Lumbantobing, E., & Chamidah, N. (2021). Implementasi Seleksi Fitur Pada Algoritma Klasifikasi Machine Learning Untuk Prediksi Penghasilan Pada Adult Income Dataset. *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA) Jakarta-Indonesia*, 2(1), 741–747.
- Fandinata, S. S., & Darmawan, R. (2020). Pengaruh Kepatuhan Minum Obat Oral Anti Diabetik Terhadap Kadar Gula Darah Pada Pasien Diabetes Mellitus Tipe II. *Jurnal Bidang Ilmu Kesehatan*, 10(1), 23–31.

<https://doi.org/10.52643/jbik.v10i1.825>

Hamid, A., & Hamdin. (2024). *Hidup Sehat Produktif Bebas Diabetes dengan Edukasi Kesehatan Tentang Penyakit Diabetes Melitus di Wilayah Kerja PKM Unit 1 Sumbawa* *Healthy Productive Life Free from Diabetes with Health Education about Diabetes Mellitus in the PKM Unit 1 Sumbawa Working*. 1(3).

Hartono, S. E. (2024). Hubungan Tingkat Pendidikan, Lama Menderita Sakit Dengan Tingkat Pengetahuan 5 Pilar Penatalaksanaan Diabetes Mellitus Di Wilayah Kerja Puskesmas Sungai Durian Kabupaten Kbu Raya Kalimantan Barat. *Journal of TSCSIKep*, 9(1), 2775–0345. <http://ejournal.annurpurwodadi.ac.id/index.php/TSCS1Kep>

Hidayat, P., Kurniawan, R., Wijaya, Y. A., & Suprapti, T. (2025). Optimasi Algoritma K-Nearest (Knn) Neighbors Pada Prediksi Risiko Penyakit Kardiovaskular. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(1). <https://doi.org/10.23960/jitet.v13i1.5864>

Kafil, M. (2019). Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 3(2), 59–66. <https://doi.org/10.36040/jati.v3i2.860>

Kementerian Kesehatan Republik Indonesia. (2025). *Review kebijakan Diabetes Melitus berbasis transformasi sistem Kesehatan dan outlook 2025*.

Lemantara, J., & Lusiani, T. (2024). Analisis Prediksi Penyakit Diabetes Pada Wanita Menggunakan Metode Naïve Bayes Dan K-Nearest Neighbor. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3). <https://doi.org/10.23960/jitet.v12i3.4911>

Maulidah, N., Supriyadi, R., Utami, D. Y., Hasan, F. N., Fauzi, A., & Christian, A. (2021). Prediksi Penyakit Diabetes Melitus Menggunakan Metode Support Vector Machine dan Naive Bayes. *Indonesian Journal on Software Engineering (IJSE)*, 7(1), 63–68. <https://doi.org/10.31294/ijse.v7i1.10279>

- Mustofa, E. E., Purwono, J., & Ludiana. (2022). Penerapan Senam Kaki Terhasap Kadar Glukosa Darah Pada Pasien Diabetes Melitus Di Wilayah Kerja Puskesmas Purwosari Kec. Metro Utara. *Jurnal Cendikia Muda*, 2(1), 78–86.
- Najah, S. (2022). LITERATUR REVIEW IMPLEMENTASI SDGs PADA KEBUTUHAN. *Jurnal Sains Edukatika Indonesia (JSEI)*, 4(1), 51–58.
- Nuraeni, F., Kurniadi, D., & Diazki, M. H. (2024). Algoritma K-Nearest Neighbor pada Kasus Dataset Imbalanced untuk Klasifikasi Kinerja Karyawan Perusahaan. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(3), 557–568. <https://doi.org/10.25126/jtiik.938144>
- Oktaviana, A., Wijaya, D. P., Pramuntadi, A., & Heksaputra, D. (2024). Prediksi Penyakit Diabetes Melitus Tipe 2 Menggunakan Algoritma K-Nearest Neighbor (K-NN). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3), 812–818. <https://doi.org/10.57152/malcom.v4i3.1268>
- Pangestu, M. S., & Fitriani, M. A. (2022). Perbandingan Perhitungan Jarak Euclidean Distance, Manhattan Distance, dan Cosine Similarity dalam Pengelompokan Data Bibit Padi Menggunakan Algoritma K-Means. *Sainteks*, 19(2), 141. <https://doi.org/10.30595/sainteks.v19i2.14495>
- Pangestu, R. A., Taslim, T., Yunefri, Y., Kursiasih, K., & Sabna, E. (2022). Optimasi Nilai k Pada Algoritma K-Nearest Neighbor Untuk Klasifikasi Pasien Covid-19 Yang Membutuhkan Ruangan ICU. *INOVTEK Polbeng - Seri Informatika*, 7(1), 147. <https://doi.org/10.35314/isi.v7i1.2481>
- Prasetyo, A. B., & Laksana, T. G. (2022). Optimasi Algoritma K-Nearest Neighbors dengan Teknik Cross Validation Dengan Streamlit (Studi Data: Penyakit Diabetes). *Journal of Applied Informatics and Computing (JAIC)*, 6(2), 194. <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Rahmadiyah, M., & Suparman, P. (2022). Penerapan Metode K-Nearest Neighbour Untuk Sistem Penentuan Peminjaman Modal Nasabah Bank Syariah Indonesia Cabang Cikarang Berbasis Website. *Jurnal Informasi Dan Komputer*, 10(2), 189–197.

- Rochman, B. F., Rahim, A., & Siswa, T. A. Y. (2024). Optimasi Algoritma KNN dengan Parameter K dan PSO Untuk Klasifikasi Status Gizi Balita. *Jurnal Media Informatika Budidarma*, 8(3), 1609. <https://doi.org/10.30865/mib.v8i3.7841>
- Romadloni, P. L., Kusuma, B. A., & Baihaqi, W. M. (2022). jurnal JATI. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), 623–628.
- Safitri, E., Rofianto, D., Purwati, N., Kurniawan, H., & Karnila, S. (2024). *Prediksi Penyakit Diabetes Melitus Menggunakan Algoritma Machine Learning Diabetes Mellitus Disease Prediction using Machine Learning Algorithms*. 12(4), 760–766. <https://doi.org/10.26418/justin.v12i4.84620>
- Siagian, Y., Hutahaean, J., Zikra Syah, A., Efendi Hutagalung, J., & Karim, A. (2024). Implementasi Metode K-Nearest Neighbours (KNN) Untuk Klasifikasi Penyakit Diabetes. *Jurnal Informatika Dan Teknologi Informasi*, 2(3), 253–262. <https://doi.org/10.56854/jt.v2i3.331>
- Surya Firmansyah, A., Aziz, A., & Ahsan, M. (2024). Optimasi K-Nearest Neighbor Menggunakan Algoritma Smote Untuk Mengatasi Imbalance Class Pada Klasifikasi Analisis Sentimen. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3341–3347. <https://doi.org/10.36040/jati.v7i6.7257>
- Syafwan, H., Siagian, F., Putri, P., Handayani, M., Tinggi Manajemen Informatika dan Komputer Royal Jln M Yamin No, S. H., & Utara, S. (2021). Forecasting Jumlah Pengangguran Di Kabupaten Asahan Menggunakan Metode Weighted Moving Average. *Jurnal Teknik Informatika Kaputama (JTIK)*, 5(2), 224–229.
- Trihardianingsih, L., & Lasatira, G. S. (2024). *Optimasi Hyperparameter GridSearchCV pada Klasifikasi Kualitas Udara menggunakan Support Vector Machine*. 1(2), 40–47.
- Yaqin, A., Kurniawan, D., & Zeniarja, J. (2025). *Optimasi Algoritma K-Nearest Neighbors Menggunakan GridSearchCV untuk Klasifikasi Penyakit Diabetes*. 16(01), 75–84. <https://doi.org/10.35970/infotekmesin.v16i1.2557>