

## Penerapan Algoritma Klasifikasi *Nearest Neighbor* Dalam Mendeteksi Penyakit Diabetes

**Siti Aqilah Sabita**

Pendidikan Matematika, FITK, Universitas Islam Negeri Sumatera Utara

Email : [sitiaqilahsabita@gmail.com](mailto:sitiaqilahsabita@gmail.com)

**Yahfizham Yahfizham**

Pendidikan Matematika, FITK, Universitas Islam Negeri Sumatera Utara

Email : [yahfizham@uinsu.ac.id](mailto:yahfizham@uinsu.ac.id)

Korespondensi penulis : [sitiaqilahsabita@gmail.com](mailto:sitiaqilahsabita@gmail.com)

**Abstract.** *The purpose of writing this article is to determine the application of the nearest neighbor classification algorithm in diabetes detection. This nearest neighbor classification algorithm is a classification method often used to classify objects based on available data. This method works by searching for the closest objects in the dataset and classifying the new objects based on the closest object category. The application of the KNN algorithm can be carried out in various fields, such as analyzing the feasibility of credit granting, classifying online news materials or diagnosing diabetes. In this article, the researcher uses a literature review research method assisted by a descriptive analysis approach, to analyze the data and by describing the data that has been previously collected where the author describes data that has been obtained from various literary sources such as journals, data and others. The data obtained will be analyzed and interpreted in accordance with the objective of this research, which is to determine the application of nearest neighbor classification to detect diabetes.*

**Keywords:** *Diabetes, Classification, K-Nearest Neighbor*

**Abstrak.** Tujuan penulisan artikel ini ialah untuk mengetahui penerapan algoritma klasifikasi nearest neighbor dalam pendeteksian penyakit diabetes. Algoritma klasifikasi nearest neighbor ini merupakan metode klasifikasi yang sering digunakan dalam mengklasifikasikan objek berdasarkan data yang tersedia. Metode ini bekerja dengan mencari objek terdekat dalam dataset dan mengklasifikasikan objek baru berdasarkan kategori objek terdekat tersebut. Penerapan algoritma KNN dapat dilakukan pada berbagai bidang, seperti analisis kelayakan pemberian kredit, pengklasifikasian dokumen berita online, dan mendiagnosis penyakit diabetes. Pada artikel ini, peneliti memakai metode penelitian studi literatur yang dibantu oleh pendekatan analisis deskriptif, guna menganalisis data dan dengan cara mendeskripsikan data yang telah terkumpul sebelumnya di mana penulis mendeskripsikan data yang telah diperoleh dari berbagai macam sumber literatur seperti jurnal, data dan lainnya. Data yang diperoleh akan dianalisis dan diinterpretasikan yang sesuai dengan tujuan penelitian ini yaitu mengetahui penerapan klasifikasi nearest neighbor untuk mendeteksi penyakit diabetes.

**Kata kunci:** Diabetes, Klasifikasi, K-Nearest Neighbor

## **LATAR BELAKANG**

Klasifikasi Nearest neighbor merupakan algoritma sederhana dalam machine learning yang sering digunakan dalam regresi dan klasifikasi. Algoritma ini mengadopsi prinsip "bird of a feather" untuk menentukan letak data baru, dengan asumsi bahwa data yang mirip akan berada dalam jarak yang dekat. Dengan memanfaatkan data-data yang ada, KNN mengklasifikasikan data-data baru yang didasari oleh kemiripan atau fungsi pada jaraknya, menempatkannya ke dalam kelas yang didominasi oleh data tetangga terdekat. Data baru kemudian menempatkannya ke dalam kelas yang didominasi oleh data tetangga terdekat berada.

Algoritma KNN dapat diterapkan dalam berbagai bidang, seperti:

1. Pemindaian medis, untuk mendeteksi penyakit berdasarkan gejala-gejala yang dialami pasien.
2. Sistem rekomendasi, untuk merekomendasikan produk atau layanan kepada pengguna berdasarkan data perilaku pengguna sebelumnya.
3. Klasifikasi citra, untuk mengklasifikasikan citra berdasarkan objek-objek yang terkandung di dalamnya.

Algoritma KNN dapat digunakan untuk mengklasifikasikan pasien baru berdasarkan datanya. Misalnya, jika data pasien baru memiliki berat badan, tinggi badan, kadar gula darah dan tekanan darah yang mirip dengan pasien-pasien diabetes, maka algoritma KNN akan mengklasifikasikan pasien tersebut sebagai pasien diabetes.

Adapun tujuan penelitian ini ialah untuk mendeskripsikan bagaimana penerapan algoritma klasifikasi nearest neighbor pada penyakit diabetes. Algoritma Nearest Neighbor telah diterapkan dalam beberapa penelitian untuk mendeteksi diabetes. Suatu penelitian menggunakan algoritma K-Nearest Neighbor digunakan dalam mengembangkan sistem pendukung keputusan untuk mengidentifikasi penyakit diabetes mellitus menggunakan faktor-faktor seperti jumlah kehamilan, kadar glukosa, tekanan darah, ketebalan lipatan kulit trisep, insulin, indeks massa tubuh (BMI), keluarga, riwayat diabetes melitus, dan usia. Studi ini mencapaitingkat keakuratan hingga 96%. Penelitian lain menggunakan algoritma K-Nearest Neighbor untuk memprediksi diabetes tipe 2. Kerangka kerja baru berbasis tetangga terdekat diusulkan dalam penelitian ketiga untuk mengklasifikasikan dua kumpulan data diabetes. Terakhir, sistem pemantauan dan konsultasi untuk manajemen pasien diabetes diusulkan dalam penelitian keempat menggunakan algoritma klasifikasi nearest neighbor.

Algoritma klasifikasi nearest neighbor (KNN) dapat diterapkan dalam mendeteksi diabetes sebagai metode klasifikasi terawasi. Ini telah diterapkan untuk mengklasifikasikan pasien ke dalam kelompok risiko tinggi dan rendah untuk mengembangkan diabetes tipe 2, dengan penilaian kinerja menggunakan metrik seperti akurasi, presisi, perolehan, dan skor F1. Algoritme ini bekerja dengan menggunakan data pelatihan untuk memprediksi apakah seseorang menderita diabetes berdasarkan faktor-faktor seperti jumlah kehamilan, kadar glukosa, tekanan darah, ketebalan lipatan kulit trisep, insulin, indeks massa tubuh (BMI), riwayat diabetes keluarga, dan usia.

Selain itu, sebuah penelitian melakukan simulasi deteksi dini pada penyakit diabetes mellitus menggunakan algoritma klasifikasi nearest Neighbors. Algoritme KNN dapat diimplementasikan ketika memprediksi apakah seseorang menderita diabetes dengan menganalisis kedekatannya dengan kumpulan data yang ada.

## **KAJIAN TEORITIS**

Algoritma klasifikasi nearest neighbor ini termasuk dalam kategori machine learning algorithm yang memiliki sifat non-parametric yang menggunakan pendekatan lazy learning. Sebagai metode dalam non-parametric, KNN tidak mengandalkan asumsi apapun mengenai distribusi data yang digunakan, sehingga tidak ditemukan jumlah parameter atau perkiraan parameter dalam model yang bersifat tetap, baik ketika data tersebut memiliki ukuran yang kecil maupun ukuran yang besar.

Algoritma K-Nearest Neighbor (KNN) bertujuan untuk mengklasifikasikan objek baru berdasarkan atribut dan juga sampel pelatihan. Pengklasifikasian dilakukan tanpa memakai model yang harus disamakan, hanya bergantung pada memori. Ketika dipakai titik query, algoritma menemukan beberapa  $k$  objek (titik pelatihan) yang jaraknya paling dekat dengan titik query. Pengklasifikasian dilakukan dengan memilih klasifikasi berdasarkan suara yang paling banyak muncul di antara  $k$  objek tersebut.

Algoritma klasifikasi nearest neighbor (KNN) memanfaatkan klasifikasi ketetanggaan sebagai nilai prediksi untuk instance query yang baru. Pendekatan KNN ini sangatlah dinilai sederhana, dengan menghitung jarak yang paling pendek dari instance query sampai sampel pelatihan sehingga dapat ditentukan ketetanggaannya. Pemilihan nilai  $k$  yang optimal bergantung pada karakteristik data.

Umumnya, pada nilai  $k$  yang tinggi dapat membuat efek noise dalam klasifikasi berkurang, tapi juga dapat menghasilkan batas antar klasifikasi tersebut semakin kurang jelas. Penentuan nilai  $k$  yang efektif biasanya dilakukan melalui optimasi parameter, seperti dengan

memakai metode cross-validation. Dalam beberapa kasus yang khusus di mana klasifikasi ini didasarkan pada satu data pelatihan terdekat (dapat disebut,  $k=1$ ), dinamakan sebagai algoritma nearest neighbor.

Keunggulan Klasifikasi Nearest Neighbor (KNN) mencakup:

1. Kemudahan implementasi. Dikarenakan sederhananya dan tingkat keakuratannya, algoritma KNN direkomendasikan sebagai suatu pengklasifikasian awal yang harusnya dipelajari terlebih dahulu oleh data scientist pemula
2. Kemudahan adaptasi. Ketika training sample baru ditambahkan, algoritma KNN secara dinamis mengadaptasi diri guna mempertimbangkan data baru, dikarenakan seluruh data pelatihan yang ada ditempatkan di memori.
3. Jumlah hyperparameter yang sedikit. KNN hanya memerlukan nilai  $k$  dan metric jaraknya, yang jumlahnya biasanya lebih sedikit daripada algoritma machine learning yang lain.

Kelemahan dalam metode KNN (K-Nearest Neighbor) melibatkan beberapa aspek:

1. Memerlukan penentuan nilai parameter  $k$ , yaitu jumlah tetangga terdekat.
2. Kurangnya kejelasan dalam memilih jenis jarak pada tahap pelatihan.
3. Tidak ada petunjuk yang pasti tentang atribut mana yang sebaiknya digunakan untuk hasil optimal.
4. Tingginya biaya komputasi karena memerlukan perhitungan jarak dari setiap instance query terhadap seluruh sampel pelatihan.

Klasifikasi memegang peranan krusial dalam ranah data science dan machine learning. Meskipun sudah lama ada, algoritma KNN tetap menjadi pilihan yang akurat untuk mengklasifikasikan pola dan model regresi. Beberapa domain di mana KNN dapat diterapkan melibatkan:

1. Penilaian kredit: Klasifikasi nearest neighbor ini membantu dalam penentuan tingkat kredit hanya dengan membandingkan individu berdasarkan karakteristik yang sama.
2. Persetujuan dalam pinjaman: Serupa dengan penilaian kredit, KNN membantu mengidentifikasi kemungkinan kegagalan pembayaran pinjaman dengan membandingkan sifat individu.
3. Preprocessing data: Algoritma KNN berguna untuk mengisi nilai-nilai yang hilang dalam dataset melalui metode imputasi data yang memperkirakan nilai yang hilang pada sebuah data

4. Pengenalan pola: Kemampuan klasifikasi ini dalam mengenali pola berguna dalam mendeteksi keanehan dalam penggunaan kartu kredit atau perilaku pembelian pelanggan.
5. Prediksi harga saham: KNN dapat memprediksi kenaikan maupun penurunan nilai saham di masa depan yang didasari oleh data histori.
6. Sistem rekomendasi: Dengan mempertimbangkan karakteristik yang sama, klasifikasi ini digunakan didalam system rekomendasi, seperti pada platform streaming online video.
7. Visi komputer: Klasifikasi ini dipakai ketika mengklasifikasikan gambar, memungkinkan pengelompokan data pada titik yang serupa, seperti membedakan antara kucing dan anjing dalam kelas yang berbeda pada aplikasi visi komputer.

Algoritma KNN juga dapat digunakan untuk mengklasifikasikan citra baru berdasarkan datanya. Misalnya, jika citra baru memiliki fitur-fitur yang mirip dengan citra-citra wajah, maka algoritma KNN akan mengklasifikasikan citra tersebut sebagai citra wajah.

## **METODE PENELITIAN**

Penelitian ini menggunakan metode penelitian studi literatur dimana dalam hal ini dibantu adanya analisis deskriptif dari berbagai sumber. Metode ini digunakan untuk mendeskripsikan hasil penelitian terkait pengimplementasian algoritma klasifikasi nearest neighbor guna klasifikasi nearest neighbor untuk mendeteksi adanya kemungkinan penyakit diabetes.

Metode studi literatur mencakup rangkaian kegiatan terkait dengan pengumpulan data-data referensi, membaca, mencatat, maupun mengolah bahan penelitiannya. M. Nazir menggambarkan studi literatur sebagai teknik mengumpulkan data dengan cara menelaah buku, catatan, laporan maupun literature lain terkait dengan masalah yang dalam proses pemecahan.

Di sisi lain, metode penelitian deskriptif memiliki tujuan untuk menggambarkan fakta sebagaimana adanya, yaitu dengan mendeskripsikan dengan objektif objek yang sedang diteliti. Metode penelitian ini dimanfaatkan untuk menjelaskan keadaan yang terjadi saat sekarang ini. Jenis metode penelitian deskriptif juga digunakan guna menguraikan gejala serta peristiwa terkait penerapan algoritma klasifikasi nearest neighbor dalam deteksi penyakit diabetes.

## **HASIL DAN PEMBAHASAN**

Penyakit diabetes disebut juga sebagai penyakit gula, merupakan suatu kondisi jangka panjang dan dicirikan oleh peningkatan kadar pada gula darah di atas angka normal. Pada penderita diabetes, tubuh tak mampu memproses glukosa agar masuk ke dalam sel yang kemudian diubah jadi energi, sehingga terjadi penimbunan gula tambahan dalam peredaran darah. Jika diabetes tidak terkontrol dengan baik, dapat menyebabkan kerusakan pada berbagai organ dan jaringan tubuh seperti ginjal, jantung, mata, dan saraf.

Menurut laporan yang diterbitkan oleh International Diabetes Federation (IDF) di tahun 2021, Indonesia menduduki peringkat kelima di dunia dengan jumlah penderita diabetes mencapai 19,47 juta, atau setara dengan 10,6% dari total populasi Indonesia. Jumlah penderita diabetes ini mengalami peningkatan yang cukup signifikan selama rentang waktu sepuluh tahun terakhir, meningkat sebanyak 167% dari jumlah penderita pada tahun 2011 yang berjumlah 7,29 juta. IDF memproyeksikan bahwa jumlah penderita diabetes di Indonesia dapat mencapai 28,57 juta pada tahun 2045, menunjukkan peningkatan sebesar 47% dibandingkan dengan tahun 2021.

Tingginya proyeksi angka penyakit diabetes dalam beberapa tahun mendatang akan menghasilkan risiko finansial tersendiri. Oleh karena itu, asuransi diperlukan dalam pengalihan risiko finansial tersebut. Apabila kelak risiko terjadi, dalam hal ini terdiagnosis memiliki penyakit diabetes, perusahaan asuransi akan menanggung biaya-biaya terkait risiko finansial tersebut.

Sampai saat ini, sudah terdapat beberapa perusahaan asuransi (Insurer) yang menyertakan penyakit diabetes sebagai salah satu manfaat pertanggung jawaban produk asuransi mereka. Insurer sebagai penanggung risiko finansial tersebut memerlukan sebuah kepastian apakah Tertanggung (Insured) benar-benar tergolong dalam penderita penyakit diabetes atau tidak. Peran dokter sangat diperlukan untuk mendiagnosis hal tersebut. Namun, jumlah dokter yang tidak sebanding dengan jumlah penderita diabetes menjadi keterbatasan tersendiri. Oleh karena itu, terdapat alternatif teknologi memanfaatkan machine learning dengan algoritma klasifikasi nearest neighbor (KNN).

Klasifikasi nearest neighbor adalah salah satu dari banyaknya metode yang dapat digunakan untuk mengklasifikasikan objek baru yang didasarkan pada sejumlah K tetangga terdekat. Algoritma KNN relatif sederhana dan mudah dipahami sehingga cukup umum untuk digunakan. Pada penerapan algoritma ini, pengklasifikasian terhadap sebuah gambar berdasarkan jarak terdekat dengan tetangganya. Hasil nilai pada jarak ini selanjutnya akan digunakan untuk melihat nilai kemiripan antara data yang diuji dan data latih.

Penelitian ini pernah dilakukan oleh Fansuari dkk pada tahun 2022 dan dipublikasikan dalam bentuk artikel yang berjudul "*Penerapan Algoritma K-Nearest Neighbor (KNN) untuk Klasifikasi Penyakit Diabetes Melitus Studi Kasus: Warga Desa Jatitengah*". Pada artikel tersebut, Fansuari membahas penelitiannya dalam menggunakan algoritma KNN terhadap dataset penyakit diabetes melitus warga desa Jatitengah. Penelitian menggunakan delapan faktor gejala yang meliputi usia, kehausan yang berlebihan, penurunan berat badan, tekanan darah tinggi, riwayat penyakit diabetes, luka yang sulit sembuh, sering buang air kecil di malam hari, dan hasil pemeriksaan gula darah. Dalam penelitian ini memanfaatkan 81 data untuk tahap pelatihan (training) dan juga 54 data untuk tahap pengujian (testing) yang telah melalui tahap pembersihan dan normalisasi. Hasil kesimpulan dari penelitian menunjukkan bahwa dengan menggunakan nilai ( $K=9$ ) pada algoritma klasifikasi nearest neighbor, didapatkan hasil klasifikasi yang menunjukkan 4 orang positif dan 50 orang negatif mengenai penyakit diabetes melitus. Penggunaan confusion matrix dalam menguji algoritma KNN menunjukkan tingkat akurasi yang tinggi mencapai 93%, presisi yang didapat mencapai 100%, recall yang didapat mencapai 60% dan FI-Score mencapai 75%. Ini menunjukkan tingkat keakuratan mencapai 93%. Berdasarkan saran dari penelitian tersebut untuk memperbanyak dataset yang digunakan, peneliti telah melakukan penerapan algoritma KNN dengan menggunakan dataset yang lebih banyak yaitu sebesar 767 sampel. Selain itu, peneliti juga menggunakan 9 atribut, dimana 8 atribut merupakan input dan 1 atribut merupakan target. Atribut yang nantinya akan digunakan sebagai landasan dalam melakukan analisis. Variabel-variabel tersebut adalah Pregnancies, Glucose, Blood, Preassure, Skin Thickness, Insulin, BMI, Diabetes Pedigree Function, dan Age. Dataset yang kemudian digunakan adalah "Diabetes Classification" yang diperoleh dari open source Kaggle.

Adapun dalam pengimplementasian Algoritma KNN digunakan sebagai pendeteksian penyakit diabetes, Contohnya terdapat data pasien yang terdiri dari fitur-fitur seperti Usia, Jenis kelamin, Riwayat keluarga, Gaya hidup dan Hasil tes medis. Data ini terdiri dari dua kategori, yaitu pasien dengan diabetes dan pasien tanpa diabetes.

1. Kumpulkan data. Data yang dibutuhkan untuk menerapkan algoritma ini adalah data pasien diabetes, baik yang sudah terdiagnosis maupun yang belum terdiagnosis. Data pasien diabetes tersebut harus mencakup fitur-fitur yang disebutkan di atas.
2. Bagi data yang didapat menjadi dua, Pertama kumpulan data pelatihan dan yang kedua kumpulan data uji. Kumpulan data pelatihan akan dimanfaatkan dalam melatih algoritma KNN, sementara kumpulan data uji akan dimanfaatkan dalam mengevaluasi kinerja algoritma KNN. Contohnya, membagi data menjadi dua kumpulan data dengan

perbandingan 70:30. Artinya, 70% dari data akan dimanfaatkan untuk kumpulan data pelatihan dan 30% dari data yang tersedia akan dimanfaatkan untuk kumpulan data uji.

3. Melatih algoritma KNN menggunakan kumpulan data pelatihan. Dalam hal ini, perlu menentukan nilai  $k$ . Disini nilai  $k$  merupakan jumlah dari tetangga terdekat dan akan digunakan dalam mengklasifikasikan data baru. Umumnya nilai  $k$  yang sering digunakan yaitu 5, 10, dan 20. Misalkan, menggunakan nilai  $k = 10$ . Artinya, algoritma KNN akan memilih 10 tetangga terdekat dari data baru untuk mengklasifikasikan data tersebut.
4. Evaluasi kinerja algoritma KNN menggunakan kumpulan data uji. Kinerja algoritma KNN dapat dievaluasi dengan menggunakan beberapa metrik, seperti akurasi, presisi, dan recall. Misalkan, akurasi algoritma KNN adalah 90%. Artinya, algoritma KNN dapat mengklasifikasikan 90% dari data uji dengan benar.

Secara rinci, langkah-langkah yang dapat dilakukan diabetes antara lain:

1. Kumpulkan data pasien yang telah didiagnosis menderita diabetes atau tidak. Dalam memperoleh data ini dapat diperoleh dari beberapa sumber, contohnya rumah sakit, klinik, atau lembaga penelitian.
2. Lakukan pra-pemrosesan data untuk membersihkan data dari kesalahan dan outlier.
3. Bagi data pasien menjadi dua kategori, yaitu pasien dengan diabetes dan pasien tanpa diabetes.
4. Pilih nilai  $k$  untuk algoritma KNN.
5. Latih algoritma KNN menggunakan kumpulan data pelatihan.
6. Evaluasi kinerja algoritma KNN menggunakan kumpulan data uji.

Adapun kelebihan dan kekurangan pada metode penerapan Algoritma Klasifikasi Nearest Neighbor antara lain

Kelebihan:

1. Algoritma ini relatif sederhana dan mudah untuk diterapkan. Algoritma nearest neighbor tidak memerlukan pemahaman yang mendalam tentang statistik atau machine learning untuk diterapkan. Algoritma ini juga dapat diimplementasikan dengan menggunakan berbagai bahasa pemrograman.
2. Algoritma ini dapat menghasilkan akurasi yang tinggi. Algoritma nearest neighbor telah terbukti dapat menghasilkan akurasi yang tinggi untuk mendeteksi penyakit diabetes. Akurasi algoritma ini dapat mencapai 90% atau lebih.



3. Algoritma ini sering sekali digunakan dalam pendeteksian penyakit diabetes dengan berbagai karakteristik, termasuk jenis diabetes, tingkat keparahan diabetes, dan komplikasi yang disebabkan oleh diabetes.

**Kekurangan:**

1. Algoritma ini dapat dipengaruhi oleh ukuran data training. Akurasi algoritma nearest neighbor dapat dipengaruhi oleh ukuran data training. Jika ukuran data training terlalu kecil, algoritma ini dapat menghasilkan akurasi yang rendah.
2. Algoritma ini dapat menghasilkan akurasi yang rendah jika data training tidak cukup representatif. Jika data training tidak cukup representatif, algoritma nearest neighbor dapat menghasilkan akurasi yang rendah. Data training yang representatif adalah data yang mencakup berbagai karakteristik dari objek yang akan diklasifikasikan.

Adapun beberapa tips untuk meningkatkan akurasi algoritma KNN dalam pendeteksian penyakit diabetes antara lain:

1. Gunakan ukuran data training yang cukup besar.
2. Pastikan data training mencakup berbagai karakteristik dari objek yang akan diklasifikasikan.
3. Gunakan nilai  $k$  yang dirasa sesuai. Yang dimana nilai  $K$  ini adalah keseluruhan jumlah objek terdekat dan sering digunakan dalam pengklasifikasian objek baru. Nilai  $K$  yang umum digunakan adalah 3, 5, atau 7.

## **KESIMPULAN DAN SARAN**

Berdasarkan hasil dari penelitian ini, peneliti dapat mengambil kesimpulan bahwa algoritma klasifikasi nearest neighbor (KNN) merupakan algoritma klasifikasi yang metodenya cukup sederhana namun efektif. Algoritma ini bekerja dengan cara mencari data latih mana yang paling mirip dengan data uji, kemudian menggunakan label data latih tersebut untuk mengklasifikasikan data uji. Algoritma ini dapat diterapkan dalam berbagai bidang, termasuk mendeteksi penyakit diabetes. Tujuan dari algoritma nearest neighbor ini ialah untuk pengklasifikasian objek-objek baru yang didasarkan pada atribut dan training sample. Dalam penerapan pada penyakit diabetes, algoritma KNN dapat membantu dalam mendiagnosis penyakit diabetes dengan memanfaatkan data yang ada dan mengklasifikasikan data berdasarkan jarak kecepatan antara data yang terdekat.

Secara keseluruhan, algoritma KNN adalah suatu metode yang sering dipakai mendeteksi penyakit diabetes secara akurat. Metode ini dapat menjadi alternatif bagi metode-

metode klasifikasi lainnya, seperti metode klasifikasi decision tree dan metode klasifikasi support vector machine

## **DAFTAR REFERENSI**

- Adminlp2m. 2023. Algoritma K-Nearest Neighbors (KNN) – Pengertian dan Penerapan. Lp2m.uma.ac.id
- Argina, A. 2020. Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes. *ndonesian Journal of Data and Science*. 1(2)
- Azizah, N dkk. 2023. Penerapan Algoritma Klasifikasi K-Nearest Neighbor pada Penyakit Diabetes. Departemen Statistika FMIPA Universitas Padjadjaran.
- Fasnuari, H dkk. 2022. Penerapan Algoritma K-Nearest Neighbor (K-Nn) Untuk Klasifikasi Penyakit Diabetes Melitus Studi Kasus: Warga Desa Jatitengah. *ANTIVIRUS: Jurnal Ilmiah Teknik Informatika*. 16(2)
- M. Yusa, E. Utami, dan E. T. Luthfi. 2016. Analisis Komparatif Evaluasi Performa Algoritma Klasifikasi pada Readmisi Pasien Diabetes. *Buana Inform*. 7(4).
- R. Wahyudi, M. Orisa, and N. Vendyansyah. 2021. Penerapan Algoritma K-Nearest Neighbors Pada Klasifikasi. *J. Mhs. Tek. Inform*. 5(2)