

HALAMAN SAMPUL

**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK KLASIFIKASI
PENYAKIT KARDIOVASKULAR**

SKRIPSI

Diajukan Untuk Memenuhi Sebagian Syarat Guna
Memperoleh Gelar Sarjana Komputer (S.Kom.)
Pada Program Studi Sistem Informasi



OLEH :

HERMAWAN NUR WAHIID

NPM : 2113030047

FAKULTAS TEKNIK DAN ILMU KOMPUTER (FTIK)
UNIVERSITAS NUSANTARA PERSATUAN GURU REPUBLIK INDONESIA

UN PGRI KEDIRI

2025

Skripsi oleh:

HERMAWAN NUR WAHID

NPM : 2113030047

Judul:

**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK KLASIFIKASI
PENYAKIT KARDIOVASKULAR**

Telah Disetujui untuk diajukan kepada

Panitia Ujian/Sidang Skripsi Program Studi Sistem Informasi

Fakultas Teknik dan Ilmu Komputer Universitas Nusantara PGRI Kediri

Tanggal: 02 Juli 2025

Pembimbing I



Rina Firliana, M.Kom
NIDN. 0731087703

Pembimbing II



Arie Nugroho, M.Kom
NIDN. 0712108103

HALAMAN PENGESAHAN

Skripsi oleh:

HERMAWAN NUR WAHIID

NPM : 2113030047

Judul:

**PENERAPAN ALGORITMA NAÏVE BAYES UNTUK KLASIFIKASI
PENYAKIT KARDIOVASKULAR**

Telah Disetujui untuk diajukan kepada

Panitia Ujian/Sidang Skripsi Program Studi Sistem Informasi

Fakultas Teknik dan Ilmu Komputer Universitas Nusantara PGRI Kediri

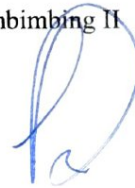
Tanggal: 02 Juli 2025

Pembimbing I



Rina Firliana, M.Kom
NIDN. 0731087703

Pembimbing II



Arie Nugroho, M.Kom
NIDN. 0712108103

HALAMAN PERSETUJUAN

Skripsi oleh:

HERMAWAN NUR WAHID

NPM: 2113030047

Judul:

**PENERAPAN ALGORITMA *NAÏVE BAYES* UNTUK KLASIFIKASI
PENYAKIT KARDIOVASKULAR**

Telah dipertahankan di depan Panitia Ujian/Sidang Skripsi

Program Studi Sistem Informasi FTIK UN PGRI Kediri

Tanggal: 09 Juli 2025

Dan Dinyatakan telah Memenuhi Persyaratan

Panitia Penguji :

1. Ketua : Rina Firliana, S.Kom, M.Kom

[.....]

2. Penguji 1 : Rini Indriati, S.Kom, M.Kom

[.....]

3. Penguji 2 : Arie Nugroho, S.Kom, M.M, M.Kom

[.....]



Mengetahui,

Dekan Fakultas Teknik dan Ilmu Komputer

Dr. Sulistiono, M.Si

NIDN. 0007076801

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini saya,

Nama : Hermawan Nur Wahiid
Jenis Kelamin : Laki-laki
Tempat/ Tanggal Lahir : Nganjuk / 20 November 2002
NPM : 2113030047
Fak/ Prodi : Teknik dan Ilmu Komputer/ Sistem Informasi

Menyatakan dengan sebenarnya, bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi, dan sepanjang pengetahuan saya tidak terdapat karya tulis atau pendapat yang pernah diterbitkan oleh orang lain, kecuali yang secara sengaja dan tertulis diacu dalam naskah ini disebutkan dalam Daftar Pustaka.

Kediri, 02 Juli 2025

Yang menyatakan

A handwritten signature in black ink is written over a red rectangular stamp. The stamp contains the text 'METERAI TEMPEL' and a unique code 'CAMX305505508' at the bottom. To the left of the stamp is a vertical black bar with the number '1000' printed vertically.

Hermawan Nur Wahiid

NPM: 2113030047

MOTTO

“Kamu harus mengakui dirimu sendiri dari dalam hati, kamu harus yakin percaya diri sebelum kamu melakukan sesuatu”

(Ningning Aespa)

ABSTRAK

Hermawan Nur Wahiid : Penerapan Algoritma *Naïve Bayes* Untuk Klasifikasi Penyakit Kardiovaskular, Skripsi, Sistem Informasi, Fakultas Teknik dan Ilmu Komputer UN PGRI Kediri, 2025.

Kata Kunci : Penyakit Kardiovaskular, *Naïve Bayes*, Klasifikasi, Data Mining, Seleksi Fitur, *SMOTE*, CRISP-DM

Penyakit kardiovaskular merupakan salah satu penyebab kematian tertinggi di dunia yang sering kali tidak terdeteksi secara dini karena keterbatasan tenaga medis, biaya pemeriksaan yang mahal, serta keterlambatan dalam proses diagnosis. Seiring dengan kemajuan teknologi informasi, pemanfaatan metode klasifikasi dalam data mining menjadi solusi yang potensial untuk mempercepat proses diagnosis awal penyakit kardiovaskular secara efisien dan akurat.

Penelitian ini bertujuan untuk menerapkan algoritma *Naïve Bayes* dalam membangun model klasifikasi penyakit kardiovaskular berbasis data rekam medis. Dataset yang digunakan berasal dari *UC Irvine Machine Learning Repository* yang terdiri atas 299 data pasien. Proses penelitian mengikuti metode CRISP-DM dan mencakup tahapan *preprocessing* seperti seleksi fitur berbasis korelasi serta penyeimbangan data menggunakan *SMOTE* untuk mengatasi permasalahan ketidakseimbangan kelas. Evaluasi model dilakukan menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*.

Hasil penelitian menunjukkan bahwa model *Naïve Bayes* yang telah diterapkan mampu mencapai akurasi sebesar 81,66%, presisi 80,43%, *recall* 74%, dan *F1-score* 77,08%. Model menunjukkan performa yang cukup baik dalam mengklasifikasikan pasien berdasarkan risiko kematian akibat penyakit kardiovaskular, terutama setelah diterapkannya *preprocessing* data yang tepat seperti seleksi fitur dan teknik *SMOTE*.

Dari hasil yang diperoleh, dapat disimpulkan bahwa algoritma *Naïve Bayes* efektif dalam membantu klasifikasi penyakit kardiovaskular dengan akurasi yang lebih tinggi dibandingkan penelitian sebelumnya. Namun, penelitian ini masih memiliki keterbatasan karena hanya menggunakan algoritma *naïve bayes* dan dataset tunggal. Untuk pengembangan selanjutnya, disarankan agar penelitian ini dikembangkan dengan membandingkan algoritma lain serta menggunakan dataset yang lebih beragam untuk meningkatkan generalisasi model dalam praktik klinis.

KATA PENGANTAR

Puji Syukur Kami panjatkan Kehadirat Allah Tuhan Yang Maha Kuasa, karena hanya atas perkenan-Nya penyusunan skripsi ini dapat diselesaikan. Skripsi dengan judul “penerapan algoritma *naïve bayes* untuk klasifikasi penyakit kardiovaskular” ini ditulis guna memenuhi sebagian syarat untuk memperoleh gelar Sarjana Komputer, pada Program Studi Sistem Informasi Fakultas Teknik dan Ilmu Komputer UN PGRI Kediri. Pada kesempatan ini diucapkan terima kasih dan penghargaan yang setulus-tulusnya kepada:

1. Dr. Zainal Afandi, M.Pd. selaku Rektor Universitas Nusantara PGRI Kediri.
2. Dr. Sulistiono, M.Si. selaku Dekan Fakultas Teknik dan Ilmu Komputer Universitas Nusantara PGRI Kediri.
3. Sucipto, M.Kom. selaku Ketua Program Studi Sistem Informasi Universitas Nusantara PGRI Kediri.
4. Rina Firliana, M.Kom. dan Arie Nugroho, M.Kom. selaku dosen pembimbing yang telah memberikan arahan, bimbingan, dan motivasi selama penyusunan skripsi ini.
5. Bapak Misiran selaku orang tua yang telah memberikan dukungan dalam do’a, dana dan motivasi dalam penyelesaian skripsi ini.
6. Ibu Umini selaku orang tua yang telah memberikan dukungan dalam do’a, dana dan motivasi dalam penyelesaian skripsi ini.
7. Saudara-saudara yang telah mendukung dalam do’a dan motivasi dalam penyelesaian skripsi ini.
8. Teman-teman yang selalu membantu memberi semangat dalam penyelesaian skripsi ini.
9. Ucapan terima kasih juga disampaikan kepada pihak-pihak lain yang tidak dapat disebutkan satu persatu, yang telah banyak membantu menyelesaikan Skripsi ini.

Disadari bahwa skripsi ini masih banyak kekurangan, maka diharapkan tegur sapa, kritik dan saran-saran, dari berbagai pihak sangat diharapkan. Akhirnya,

disertai harapan semoga skripsi ini ada manfaatnya bagi kita semua, khususnya bagi dunia pendidikan, meskipun hanya ibarat setitik air bagi samudra yang luas.

Kediri, 02 Juli 2025

HERMAWAN NUR WAHID
NPM : 2113030047

DAFTAR ISI

Halaman Sampul.....	i
Halaman Persetujuan.....	ii
Halaman Pengesahan.....	iii
Halaman Pernyataan.....	iv
Halaman Motto.....	v
ABSTRAK	vi
KATA PENGANTAR	vii
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah.....	1
B. Identifikasi Masalah.....	2
C. Batasan Masalah.....	2
D. Rumusan Masalah.....	3
E. Tujuan Penelitian.....	3
F. Manfaat Penelitian.....	3
BAB II KAJIAN PUSTAKA	4
A. Kajian Teori.....	4
1. Penyakit Kardiovaskular.....	4
2. Data Mining.....	4
3. <i>Naïve Bayes</i>	5
4. <i>Jupyter Notebook</i>	5
5. <i>Preprocessing</i>	6
6. Matriks Evaluasi dalam Klasifikasi.....	6
7. <i>SMOTE</i>	7
8. Seleksi Fitur.....	8

9. Metode CRISP-DM.....	8
B. Kajian Hasil Penelitian Terdahulu	9
BAB III METODE PENELITIAN	13
A. <i>Business Understanding</i> (Pemahaman Bisnis).....	13
B. <i>Data Understanding</i> (Pemahaman Data).....	14
C. <i>Data Preparation</i> (Persiapan Data).....	14
1. Pemisahan Fitur dan Target.....	15
2. Analisis Korelasi.....	15
3. Seleksi Fitur.....	15
4. Pembagian Data.....	16
5. Penyeimbangan Data dengan <i>SMOTE</i>	16
D. <i>Modeling</i> (Pemodelan).....	16
E. <i>Evaluation</i> (Evaluasi).....	17
F. <i>Deployment</i> (Penyebaran).....	17
BAB IV HASIL DAN PEMBAHASAN.....	18
A. <i>Business Understanding</i>	18
B. <i>Data Understanding</i>	19
C. <i>Data Preparation</i>	22
D. <i>Modeling</i>	24
1. Memuat Dataset.....	24
2. Memisahkan Fitur dan Target.....	26
3. Analisis Korelasi.....	26
4. Seleksi Fitur.....	28
5. Pembagian Data.....	29
6. Cek Keseimbangan Data.....	29
7. Menyeimbangkan Data dengan <i>SMOTE</i>	31
8. Cek Ulang Distribusi Setelah <i>SMOTE</i>	31
9. Menerapkan Model <i>Naïve Bayes</i>	33
10. Evaluasi Model.....	34
11. <i>Confusion Matrix</i>	36
12. Menampilkan Hasil Evaluasi.....	36
E. <i>Evaluation</i>	38
1. Evaluasi Berdasarkan Proporsi Data <i>Training</i> dan <i>Testing</i>	39

2. Evaluasi Tanpa Seleksi Fitur	40
3. Evaluasi Tanpa <i>SMOTE</i>	42
4. Hasil Evaluasi Keseluruhan	42
F. <i>Deployment</i>	43
BAB V PENUTUP	44
A. Kesimpulan	44
B. Saran	44
DAFTAR PUSTAKA	46
LAMPIRAN	50

DAFTAR GAMBAR

Gambar 2. 1 Distribusi Kelas tidak Seimbang Sebelum dilakukan <i>SMOTE</i>	7
Gambar 2. 2 Kerangka Berpikir	11
Gambar 3. 1 Alur CRISP-DM.....	13
Gambar 4. 1 Kode Program Memuat Dataset	25
Gambar 4. 2 Kode Program Memisahkan Fitur dan Target.....	26
Gambar 4. 3 Kode Program Analisis Korelasi.....	27
Gambar 4. 4 Tampilan Hasil Korelasi Fitur Terhadap Target <i>Death_Event</i>	27
Gambar 4. 5 Kode Program Seleksi Fitur	29
Gambar 4. 6 Kode Program Pembagian Data	29
Gambar 4. 7 Kode Program Cek Keseimbangan Data.....	29
Gambar 4. 8 Distribusi Kelas (Target) Sebelum <i>SMOTE</i>	30
Gambar 4. 9 Kode Program Menyeimbangkan Data dengan <i>SMOTE</i>	31
Gambar 4. 10 Kode Program Cek Ulang Distribusi Setelah <i>SMOTE</i>	32
Gambar 4. 11 Distribusi Kelas (Target) Setelah <i>SMOTE</i>	33
Gambar 4. 12 Kode Program Menerapkan Model <i>Naive Bayes</i>	34
Gambar 4. 13 Kode Program Evaluasi Model.....	35
Gambar 4. 14 Hasil Evaluasi Model.....	35
Gambar 4. 15 Kode Program <i>Confusion Matrix</i>	36
Gambar 4. 16 Kode Program Menampilkan Hasil Evaluasi	36
Gambar 4. 17 <i>Confusion Matrix</i>	37

DAFTAR TABEL

Tabel 4.1 Tampilan Dataset.....	26
Tabel 4.2 Hasil Fitur Prioritas Untuk Korelasi Terhadap Target.....	28
Tabel 4.3 Distribusi kelas sebelum <i>SMOTE</i>	30
Tabel 4.4 Distribusi kelas setelah <i>SMOTE</i>	32
Tabel 4.5 Hasil Evaluasi Berdasarkan Proporsi Data <i>Training</i> dan <i>Testing</i>	39
Tabel 4.6 Hasil Evaluasi Tanpa Seleksi Fitur.....	41
Tabel 4.7 Metriks Evaluasi Model.....	43

DAFTAR LAMPIRAN

Lampiran 1 Kartu Bimbingan Skripsi.....	50
Lampiran 2 Surat Keterangan Bebas <i>Similarity</i>	51
Lampiran 3 Bukti Halaman Awal cek <i>Similarity</i>	52
Lampiran 4 Bukti Submit Artikel/Loa.....	53
Lampiran 5 Berita Acara Ujian SIAKAD.....	54
Lampiran 6 Lembar Revisi Ujian SIAKAD.....	55

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Penyakit kardiovaskular adalah gangguan kesehatan yang terjadi akibat adanya masalah pada fungsi jantung dan pembuluh darah, seperti hipertensi, penyakit jantung koroner, stroke, serta gagal jantung (Nugroho et al., 2020). Penyakit ini juga berpotensi menyebabkan kecacatan serta menurunkan kualitas hidup penderita terkait kesehatan. Penyakit kardiovaskular merupakan salah satu penyebab utama kematian di seluruh dunia. Menurut data Organisasi Kesehatan Dunia (WHO), lebih dari 17 juta orang meninggal setiap tahun akibat penyakit yang berkaitan dengan jantung dan pembuluh darah. Faktor risiko penyakit kardiovaskular meliputi beberapa aspek, salah satunya adalah usia; semakin tua seseorang, semakin besar risikonya terkena penyakit ini. Faktor lainnya termasuk riwayat keluarga dengan penyakit kardiovaskular, diabetes, tekanan darah tinggi, obesitas, gaya hidup yang tidak sehat, serta stres. Dengan prevalensi yang terus meningkat, deteksi dini penyakit kardiovaskular menjadi sangat penting untuk mengurangi angka kematian dan meningkatkan kualitas hidup penderita.

Kemajuan teknologi informasi dalam mengolah data dalam jumlah besar selalu erat kaitannya dengan perkembangan data mining. Data mining adalah suatu proses yang digunakan pada basis data besar dengan memanfaatkan kecerdasan buatan, teknik matematika, statistik, dan *machine learning* untuk menarik kesimpulan serta mengidentifikasi informasi dan pengetahuan yang berharga dan relevan (Priyanto et al., 2024). Oleh sebab itu, berbagai algoritma telah dikembangkan untuk menangani data berukuran besar, salah satunya adalah algoritma klasifikasi (Sholekhah et al., 2024). Klasifikasi dalam data mining adalah proses menempatkan objek-objek ke dalam salah satu kategori yang telah ditentukan sebelumnya (Desiani et al., 2022a). Metode ini bertujuan untuk menentukan nilai dari suatu variabel yang belum diketahui berdasarkan variabel lain yang sudah diketahui sebelumnya (Riany & Testiana, 2023). Algoritma klasifikasi dapat membantu memprediksi kemungkinan seseorang menderita penyakit kardiovaskular berdasarkan data historis atau riwayat medis. Salah satu

algoritma yang sering digunakan dalam klasifikasi data medis adalah *Naïve Bayes*.

Naïve Bayes merupakan metode statistik yang sederhana namun mampu memberikan akurasi yang tinggi dan tingkat kesalahan yang rendah dalam proses klasifikasi (Dewi et al., 2024). *Naïve Bayes Classifier*, yang didasarkan pada Teorema *Bayes*, memprediksi kemungkinan di masa depan dengan memanfaatkan pengalaman sebelumnya melalui pendekatan statistik dan probabilitas (Wardhani et al., 2024). *Naïve Bayes* adalah metode yang menggunakan dua pendekatan interpretasi yang sedikit berbeda. Berdasarkan pandangan *Bayesian*, model ini menggambarkan sejauh mana tingkat keyakinan subjektif perlu berubah secara rasional ketika mendapatkan informasi atau bukti baru. Sementara itu, menurut pandangan frekuentis, model ini merepresentasikan probabilitas terbalik yang dianalisis melalui dua skenario kejadian (Ari Bianto et al., 2019).

B. Identifikasi Masalah

Penyakit kardiovaskular menjadi salah satu penyebab utama kematian di dunia, sehingga diagnosis dini menjadi hal yang sangat penting untuk mencegah komplikasi lebih lanjut serta meningkatkan kualitas hidup pasien. Namun, proses diagnosis yang akurat seringkali membutuhkan waktu lama, biaya yang mahal, serta keterlibatan tenaga medis dengan keahlian khusus. Tantangan ini menimbulkan kebutuhan akan metode yang dapat membantu para tenaga medis dalam melakukan diagnosis awal dengan lebih cepat dan akurat. Dalam bidang klasifikasi data medis, algoritma *Naïve Bayes* telah banyak digunakan karena kemampuannya dalam memproses data dengan cepat dan menghasilkan prediksi yang cukup akurat. Oleh karena itu, penerapan algoritma ini diharapkan dapat memberikan alternatif yang lebih efisien untuk mengklasifikasikan penyakit kardiovaskular secara lebih cepat dan akurat.

C. Batasan Masalah

1. Data yang digunakan untuk penelitian ini adalah dataset penyakit jantung kardiovaskular yang telah tersedia secara publik yang diakses melalui *UC Irvine Machine Learning Repository* dengan alamat website berikut ini: <https://archive.ics.uci.edu/dataset/519/heart+failure+clinical+records>.

2. Penelitian ini hanya akan menggunakan algoritma *Naïve Bayes* untuk klasifikasi penyakit kardiovaskular.
3. Parameter evaluasi yang digunakan untuk perbandingan performa algoritma adalah akurasi, presisi, *recall*, dan *f1-score*.
4. Penelitian ini hanya berfokus pada klasifikasi penyakit kardiovaskular dan tidak mencakup prediksi penyakit lainnya.
5. Penelitian ini hanya menggunakan bahasa pemrograman *python*.
6. Dataset penelitian ini berjenis teks dengan format tabel.

D. Rumusan Masalah

Berdasarkan uraian latar belakang dan identifikasi masalah, maka dapat dirumuskan permasalahan dalam penelitian ini sebagai berikut: Bagaimana penerapan algoritma *naïve bayes* untuk klasifikasi penyakit kardiovaskular?

E. Tujuan Penelitian

Pemaparan Penelitian ini bertujuan untuk:

1. Mengevaluasi performa model klasifikasi menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*.
2. Menerapkan algoritma *Naïve Bayes* untuk membangun model klasifikasi penyakit kardiovaskular.

F. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat di berbagai bidang. Bagi dunia akademik, penelitian ini akan memberikan wawasan tentang perbandingan algoritma *Naïve Bayes* dalam bidang klasifikasi penyakit, khususnya penyakit kardiovaskular. Bagi praktisi medis, hasil penelitian ini dapat menjadi referensi mengenai metode klasifikasi yang lebih efektif untuk mendukung proses diagnosis penyakit kardiovaskular, sehingga membantu dalam upaya deteksi dini dan penanganan yang lebih baik. Selain itu, bagi peneliti selanjutnya, penelitian ini dapat menjadi acuan dan dasar untuk pengembangan penelitian lebih lanjut terkait penggunaan *machine learning* dalam diagnosis penyakit atau bidang kesehatan lainnya.

DAFTAR PUSTAKA

- Abdullah Muttaqin, A., Alam, S., & Andayani Komara, M. (2024). ANALISIS SENTIMEN ISU KECURANGAN PEMILU 2024 BERDASARKAN OPINI PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE CRISP-DM DENGAN ALGORITMA NAÏVE BAYES CLASSIFIER. *Jurnal Mahasiswa Teknik Informatika (JATI)*, 8(5). <https://doi.org/https://doi.org/10.36040/jati.v8i5.10821>
- Ari Bianto, M., Kusriani, & Sudarmawan. (2019). Perancangan Sistem Klasifikasi Penyakit Jantung Menggunakan Naïve Bayes Designing a Heart Disease Classification System Using Naïve Bayes. *Citec Journal*, 6(1). <https://doi.org/https://doi.org/10.24076/citec.2019v6i1.231>
- Cholifah Sastya, N., & Nugraha, I. (2023). Penerapan Metode CRISP-DM dalam Menganalisis Data untuk Menentukan Customer Behavior di MeatSolution. *Jurnal Pendidikan Dan Aplikasi Industri*, 10. <https://doi.org/https://doi.org/10.33592/unistek.v10i2.3079>
- Depari, D. H., Widiastiwi, Y., & Santoni, M. M. (2022). Perbandingan Model Decision Tree, Naive Bayes dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung. *Informatik : Jurnal Ilmu Komputer*, 18(3), 239. <https://doi.org/10.52958/iftk.v18i3.4694>
- Desiani, A., Akbar, M., Irmeilyana, I., & Amran, A. (2022). Implementasi Algoritma Naïve Bayes dan Support Vector Machine (SVM) Pada Klasifikasi Penyakit Kardiovaskular. *Jurnal Teknik Elektro Dan Komputasi (ELKOM)*, 4(2), 207–214. <https://doi.org/10.32528/elkom.v4i2.7691>
- Devia, A., & Soewito, B. (2023). Analisis Perbandingan Metode Seleksi Fitur untuk Mendeteksi Anomali pada Dataset CIC-IDS-2018. *Jurnal Teknologi Dan Sistem Informasi Bisnis-JTEKSIS*, 5(4), 572. <https://doi.org/10.47233/jteksis.v5i4.1069>
- Dewi, A., Safira Surya, A., & Yamasari, Y. (2024). Penerapan Algoritma Naïve Bayes (NB) untuk Klasifikasi Penyakit Jantung. *Journal of Informatics and Computer Science*, 05. <https://doi.org/https://doi.org/10.26740/jinacs.v5n03.p447-455>
- Dhewayani, F. N., Amelia, D., Alifah, D. N., Sari, B. N., & Jajuli, M. (2022). Implementasi K-Means Clustering untuk Pengelompokkan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. *Jurnal Teknologi Dan Informasi*, 12. <https://doi.org/10.34010/jati.v12i1>
- Muzaki, M. N., Firliana, R., Indriati, R., Wardani, A. S., & Daniati, E. (2025). Multilevel Enkripsi Menggunakan Kombinasi Algoritma Kriptografi Base64 dan Pepper. *Jurnal Qua Teknika*, 15(1), 14–24. <https://doi.org/https://doi.org/10.35457/quateknika.v15i01.4364>
- Navisa, S., Hakim, L., & Nabilah, A. (2021). Komparasi Algoritma Klasifikasi Genre Musik pada Spotify Menggunakan CRISP-DM. *Jurnal Sistem Cerdas*, 04.

<https://doi.org/https://doi.org/10.37396/jsc.v4i2.162>

- Nugroho, A., Bimo Gumelar, A., Sooi, A. G., Sarvasti, D., & Tahalele, P. L. (2020). Perbandingan Performansi Algoritma Pengklasifikasian Terpandu Untuk Kasus Penyakit Kardiovaskular. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 1(3), 998–1006. <https://doi.org/https://doi.org/10.29207/resti.v4i5.2316>
- Nugroho, A., & Harini, D. (2024). Teknik Random Forest untuk Meningkatkan Akurasi Data Tidak Seimbang. *JSITIK*, 2(2). <https://doi.org/10.53624/jsitik.v2i2.XX>
- Nurjanah, I., Karaman, J., Widaningrum, I., Mustikasari, D., & Sucipto. (2023). Penggunaan Algoritma Naive Bayes Untuk Menentukan Pemberian Kredit Pada Koperasi Desa. In *Journal of Computer Science and Information Technology E-ISSN* (Vol. 3, Issue 2). <https://doi.org/https://doi.org/10.47065/explorer.v3i2.766>
- Parlindungan Pane, J., Simorangkir, L., & Sari Br Saragih, P. I. (2022). FAKTOR-FAKTOR RISIKO PENYAKIT KARDIOVASKULAR BERBASIS MASYARAKAT. *Jurnal Penelitian Perawat Profesional*. <https://doi.org/https://doi.org/10.37287/jppp.v4i4.1218>
- Pramayasa, K., Md, I., Maysanjaya, D., Ayu, G., & Diatri Indradewi, A. (2023). Analisis Sentimen Program Mbkm Pada Media Sosial Twitter Menggunakan KNN Dan SMOTE. *SINTECH Journal*, 6. <https://doi.org/https://doi.org/10.31598/sintechjournal.v6i2.1372>
- Priyanto, E., Daniati, E., & Ristyawan, A. (2024). Implementasi Metode K-Nearest Neighbor Untuk Memprediksi Kondisi Cuaca Penulis Korespondensi. *SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 8, 2549–7952. <https://doi.org/https://doi.org/10.29407/inotek.v8i1.4954>
- Putra, F., Tahiyat, H. F., Ihsan, R. M., Rahmaddeni, R., & Efrizoni, L. (2024). Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(1), 273–281. <https://doi.org/10.57152/malcom.v4i1.1085>
- Putri, R. W., Ristyawan, A., & Muzaki, M. N. (2022). Comparison Performance of K-NN and NBC Algorithm for Classification of Heart Disease. *JTECS: Jurnal Sistem Telekomunikasi Elektronika Sistem Kontrol Power Sistem Dan Komputer*, 2(2), 143. <https://doi.org/10.32503/jtecs.v2i2.2708>
- Riany, A. F., & Testiana, G. (2023). PENERAPAN DATA MINING UNTUK KLASIFIKASI PENYAKIT JANTUNG KORONER MENGGUNAKAN ALGORITMA NAÏVE BAYES. *MDP Student Conference (MSC)*. <https://doi.org/https://doi.org/10.35957/mdp-sc.v2i1.4388>
- Ristyawan, A., Nugroho, A., & Amarya, T. K. (2025). Optimasi Preprocessing Model Random Forest Untuk Prediksi Stroke. *Jurnal Teknik Informatika Dan Sistem Informasi*, 12(1), 29–44. <https://doi.org/https://doi.org/10.35957/jatisi.v12i1.9587>

- Rizqullah, M. F., Raihana, N. T., & Jambak, M. I. (2024). Komparasi Penerapan Algoritma C4.5, K-Nearest Neighbor, dan Naïve Bayes untuk Keberlangsungan Pasien Gagal Jantung. *KLIK: Kajian Informatika Dan Komputer*, 4(5), 2580–2587. <https://doi.org/10.30865/klik.v4i5.1788>
- Sahar. (2020). Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Classiffier pada Data Set Penyakit Jantung. *Indonesian Journal of Data and Science (IJODAS)*, 1(3), 79–86. <https://doi.org/https://doi.org/10.33096/ijodas.v1i3.20>
- Setiawan, Y. (2023). Data Mining berbasis Nearest Neighbor dan Seleksi Fitur untuk Deteksi Kanker Payudara. *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, 8(2). <https://doi.org/https://doi.org/10.30591/jpit.v8i2.4994>
- Sholekhah, F., Putri, A. D., Rahmaddeni, R., & Efrizoni, L. (2024). Perbandingan Algoritma Naïve Bayes dan K-Nearest Neighbors untuk Klasifikasi Metabolik Sindrom. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 507–514. <https://doi.org/10.57152/malcom.v4i2.1249>
- Sidiq, S., Alfian, & Shobi Maburur, N. (2025). Pengembangan Model Prediksi Risiko Diabetes Menggunakan Pendekatan AdaBoost dan Teknik Oversampling SMOTE. *Jurnal Ilmiah Informatika Dan Ilmu Komputer*, 4. <https://doi.org/10.58602/jima-ilkom.v4i1.41>
- Singgalen, Y. A. (2023). Penerapan Metode CRISP-DM untuk Optimalisasi Strategi Pemasaran STP (Segmenting, Targeting, Positioning) Layanan Akomodasi Hotel, Homestay, dan Resort. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(4), 1980. <https://doi.org/10.30865/mib.v7i4.6896>
- Suhandha, Y., Kurniati, I., & Norma, S. (2020). Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik. *Jurnal Teknologi Informatika Dan Komputer*, 6(2), 12–20. <https://doi.org/10.37012/jtik.v6i2.299>
- Syifa, H. A. N., Nugroho, A., & Firliana, R. (2023). Perbandingan Algoritma Naïve Bayes Classifier Dan K-Nearest Neighbors Untuk Analisis Sentimen Covid-19 Di Twitter. *Jurnal Ilmiah Informatika*, 11. <https://doi.org/https://doi.org/10.33884/jif.v11i01.7069>
- Wardhani, A. F. K., Kamilatutsaniya, N., Alamsyah, M. A., Daniati, E., & Ristyawan, A. (2024). Perbandingan Algoritma Naive Bayes Dan Desicion Tree Dalam Pengujian Data Anemia Menggunakan K-Fold Cross Validation. *SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 8. <https://doi.org/https://doi.org/10.29407/inotek.v8i1.4945>
- Wibisono, A. B., & Fahrurozi, A. (2019). PERBANDINGAN ALGORITMA KLASIFIKASI DALAM PENGKLASIFIKASIAN DATA PENYAKIT JANTUNG KORONER. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 24(3), 161–170. <https://doi.org/10.35760/tr.2019.v24i3.2393>

Yudiana, Y., Yulia Agustina, A., & Khofifah, N. (2023). Prediksi Customer Churn Menggunakan Metode CRISP-DM Pada Industri Telekomunikasi Sebagai Implementasi Mempertahankan Pelanggan. *IJIEB: Indonesian Journal of Islamic and Business*, 8(1). <https://doi.org/https://doi.org/10.30631/ijiebv8i1.1710>

