

**PENERAPAN *DATA MINING* UNTUK IDENTIFIKASI RISIKO TERKENA
PENYAKIT JANTUNG MENGGUNAKAN *RANDOM FOREST***

SKRIPSI

Program Studi Sistem Informasi

Jenjang Sarjana



Oleh

Galih Ananda Dwisari

09031282126107

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS SRIWIJAYA

2025

HALAMAN PENGESAHAN

SKRIPSI

**Penerapan Data Mining Untuk Identifikasi Risiko Terkena Penyakit Jantung
Menggunakan Random Forest**

Sebagai salah satu syarat untuk penyelesaian studi di

Program Studi S1 Sistem Informasi

Oleh:

GALIH ANANDA DWISARI

09031282126107

Pembimbing 1

**: M. Husni Syahbani, M.T.
NIP. 198604202023211023**

Mengetahui

Ketua Jurusan Sistem Informasi



**Ahmad Rifai, S.T., M.T.
1979102010121003**

HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan di bawah ini:

Nama : Galih Ananda Dwisari
NIM : 09031282126107
Program Studi : Sistem Informasi Reguler
Jenjang : Strata 1 (S1)
Judul Skripsi : Penerapan *Data Mining* Untuk Identifikasi Risiko
Terkena Penyakit Jantung Menggunakan *Random Forest*

Hasil Pengecekan Software Ithenticate/Turnitin: 15%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan skripsi ini maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tanpa paksaan dari siapa pun.



Palembang, 3 Juli 2025
Penulis,



Galih Ananda Dwisari
NIM. 09031282126107

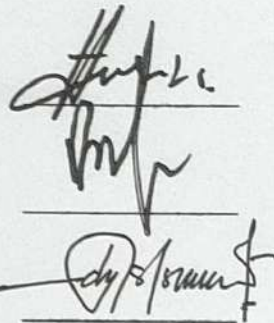
HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Jumat
Tanggal : 13 Juni 2025
Nama : Galih Ananda Dwisari
NIM : 09031282126107
Judul Tugas Akhir : Penerapan *Data Mining* Untuk Identifikasi
Risiko Terkena Penyakit Jantung
Menggunakan *Random Forest*

Komisi Penguji:

1. Pembimbing : M. Husni Syahbani, M.T.
2. Ketua : Dedy Kurniawan, M.Sc.
3. Penguji : Dinna Yunika Hardiyanti, M.T.



Mengetahui,

Ketua Jurusan Sistem Informasi



Ahmad Rifai, S.T., M.T.

NIP. 1979102010121003

HALAMAN PERSEMBAHAN

MOTTO

“Menjalani hidup dengan niat yang lurus, menebar manfaat
dengan hati yang tulus.”

Skripsi ini penulis dedikasikan kepada:

- ❖ Kedua orang tua tercinta dengan doanya yang tak pernah putus.
- ❖ Keluarga besar yang selalu memberi dukungan.
- ❖ Sahabat seperjuangan yang setia menemani.
- ❖ Para guru, atas ilmu yang bermanfaat.
- ❖ Dan untuk diri sendiri.

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Allah SWT. atas segala rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “Penerapan *Data Mining* Untuk Identifikasi Risiko Terkena Penyakit Jantung Menggunakan *Random Forest*”. sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Sriwijaya.

Penyusunan Laporan Tugas Akhir ini merupakan proses yang penuh tantangan. Namun, berkat arahan, bantuan, dukungan, semangat, dan doa dari berbagai pihak, penulis mampu melewati setiap tahapan dan menyelesaikan laporan ini dengan baik. Oleh karena itu, dengan rasa hormat yang tulus, penulis menyampaikan terima kasih kepada:

1. Kepada Allah SWT, atas limpahan rahmat, kesehatan, kesempatan, kemudahan, serta segala keberkahan yang telah diberikan kepada penulis selama proses penyusunan Skripsi ini.
2. Untuk kedua orang tua tercinta, yang selalu menjadi sumber kekuatan dan inspirasi.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si., selaku Dekan Fakultas Ilmu Komputer.
4. Bapak Ahmad Rifai, S.T., M.T., selaku Ketua Jurusan Sistem Informasi, Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Bapak Muhammad Husni Syahbani, M.T., sebagai Dosen Pembimbing Tugas Akhir.
6. Ibu Allsela Meiriza S.Kom. M.T., sebagai Dosen Pembimbing Akademik.

7. Seluruh dosen di Fakultas Ilmu Komputer, yang telah memberikan ilmu, wawasan, serta pengalaman berharga yang sangat bermanfaat dalam proses belajar penulis selama perkuliahan.
8. Tim administrasi Program Studi Sistem Informasi, atas bantuan dan kemudahan yang diberikan dalam menyelesaikan berbagai keperluan akademik dari awal hingga akhir masa studi.
9. Keluarga besar yang telah bersedia membantu memberikan dukungan baik berupa moral dan finansial kepada penulis selama menyelesaikan Laporan Tugas Akhir ini.
10. Rachel Aqilla Maghfiroh dan Imam Zada Ainullah serta,
11. Seluruh teman seperjuangan Jurusan Sistem Informasi Angkatan 2021.

Palembang, 3 Juli 2025



Galih Ananda Dwisari
NIM. 09031282126107

PENERAPAN *DATA MINING* UNTUK IDENTIFIKASI RISIKO TERKENA PENYAKIT JANTUNG MENGGUNAKAN *RANDOM FOREST*

Oleh

Galih Ananda Dwisari

09031282126107

ABSTRAK

Penyakit jantung merupakan salah satu penyebab kematian tertinggi di dunia, sehingga upaya deteksi dini terhadap risikonya sangat penting. Penelitian ini bertujuan untuk menerapkan teknik *data mining* dalam mengidentifikasi risiko terkena penyakit jantung dengan menggunakan algoritma *Random Forest*. Algoritma ini dipilih karena kemampuannya dalam menangani data yang kompleks serta memberikan akurasi tinggi dalam proses klasifikasi. Penelitian ini menghasilkan sistem informasi yang terdiri dari dua komponen utama, yaitu API yang memuat model prediksi dan menangani proses klasifikasi, serta situs web yang berfungsi sebagai antarmuka pengguna untuk memasukkan data dan melihat hasil prediksi. *Dataset* yang digunakan bersumber dari platform Kaggle. Model dilatih dengan parameter yang dioptimalkan menggunakan *Optuna*. Hasil evaluasi menunjukkan bahwa model yang telah dilatih mampu mengklasifikasikan risiko penyakit jantung dengan tingkat rata - rata akurasi sebesar 85.83%, serta memberikan informasi tambahan berupa faktor risiko yang paling berpengaruh. Sistem ini diharapkan dapat digunakan sebagai sarana edukasi untuk meningkatkan kesadaran masyarakat terhadap faktor risiko penyakit jantung.

Kata kunci: *data mining*, penyakit jantung, *Random Forest*, prediksi risiko, sistem informasi

APPLICATION OF DATA MINING FOR IDENTIFYING THE RISK OF HEART DISEASE USING RANDOM FOREST

By

Galih Ananda Dwisari

09031282126107

ABSTRACT

Heart disease is one of the leading causes of death worldwide, making early detection of its risks crucial. This research aims to apply data mining techniques to identify the risk of heart disease using the Random Forest algorithm. This algorithm was chosen for its ability to handle complex data and provide high accuracy in classification processes. This study produced an information system consisting of two main components: an API that contains the prediction model and handles the classification process, and a website that serves as a user interface for inputting data and viewing prediction results. The dataset used was sourced from the Kaggle platform. The model was trained with parameters optimized using Optuna. The evaluation results show that the trained model is capable of classifying heart disease risk with an average accuracy rate of 85.83%, as well as providing additional information on the most influential risk factors. It is hoped that this system can be used as an educational tool to increase public awareness of heart disease risk factors.

Keywords: data mining, heart disease, Random Forest, risk prediction, information system

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN BEBAS PLAGIAT	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
ABSTRAK	viii
<i>ABSTRACT</i>	ix
DAFTAR ISI	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah.....	5
1.6 Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA	8
2.1 Data Mining	8
2.2 Decision Tree	9

2.3 Random Forest	10
2.4 CRISP-DM.....	11
2.5 Sistem Informasi Berbasis Website.....	12
2.6 Penyakit Jantung	13
2.6.1 Pengertian Penyakit Jantung	13
2.6.2 Faktor Risiko Penyakit Jantung	14
2.6.3 Deteksi Dini Penyakit Jantung.....	14
2.6.4 Deteksi Dini Penyakit Jantung.....	15
2.7 Penelitian Tedahulu	15
BAB III METODOLOGI PENELITIAN.....	20
3.1 Pendekatan Penelitian	20
3.2 Metodologi Data Mining.....	21
3.2.1 Pemahaman Bisnis	22
3.2.2 Pemahaman Data.....	22
3.2.3 Persiapan Data.....	22
3.2.4 Pemodelan	23
3.2.5 Evaluasi	24
3.2.6 Deployment.....	25
BAB IV HASIL DAN PEMBAHASAN.....	29
4.1 Pemahaman Bisnis	29
4.2 Pemahaman Data.....	29

4.3 Persiapan Data.....	34
4.4 Pemodelan	35
4.5 Evaluasi	38
4.6 Deployment	39
4.6.1 Hasil Analisis Kebutuhan Sistem	40
4.6.2 Hasil Perancangan Sistem	41
4.6.3 Pengembangan dan Implementasi API.....	42
4.6.4 Pengembangan Antarmuka Pengguna.....	43
4.6.5 Hasil Pengujian	45
4.6.6 <i>Deployment</i> Menggunakan Docker.....	46
BAB V KESIMPULAN DAN SARAN.....	47
5.1 Kesimpulan	47
5.2 Saran.....	47
DAFTAR PUSTAKA	49

DAFTAR TABEL

Tabel 4.1 Sampel dataset.....	30
Tabel 4. 2 Struktur data	31
Tabel 4. 3 Sebaran nilai pada fitur numerik	31
Tabel 4. 4 Sebaran nilai pada fitur kategorikal	32
Tabel 4. 5 Hasil Optimasi <i>Hyperparameter</i>	36
Tabel 4. 6 Lima peringkat kepentingan fitur tertinggi	37
Tabel 4.7 Endpoint API.....	43
Tabel 4.8 Pengujian sistem.....	46

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian.....	20
Gambar 4. 1 Distribusi fitur target	33
Gambar 4. 2 Distribusi fitur kategorikal	33
Gambar 4. 3 Kode program evaluasi Performa Model	38
Gambar 4. 4 Hasil evaluasi performa model.....	38
Gambar 4.5 Diagram Konteks.....	41
Gambar 4.6 DFD Level 0.....	42
Gambar 4.7 Halaman Masukan.....	44
Gambar 4.8 Hasil prediksi	44
Gambar 4.9 Tampilan halaman riwayat prediksi	45
Gambar 4.10 Pengujian Postman	45

DAFTAR LAMPIRAN

Lampiran 1 Kode program data mining	A-1
Lampiran 2 Surat kesediaan membimbing.....	B-1
Lampiran 3 Surat keputusan pembimbing mahasiswa.....	C-1
Lampiran 4 Hasil pengecekan plagiarisme	D-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital di Indonesia saat ini telah membawa perubahan besar dalam berbagai aspek kehidupan masyarakat. Akses informasi menjadi lebih mudah dan cepat, serta munculnya layanan daring memungkinkan masyarakat untuk berinteraksi dan berkomunikasi secara lebih luas (Rabbani & Najicha, 2023). Perkembangan ini juga menghasilkan volume data yang sangat besar akibat aktivitas masyarakat di berbagai platform digital, termasuk media sosial, sistem informasi, dan transaksi daring (R. E. Santoso et al., 2022).

Kumpulan data yang memiliki ukuran besar dan kompleks diberi istilah *big data*, yang memiliki karakteristik volume tinggi, keragaman format, serta kompleksitas dalam pengelolaannya. *Big data* telah dimanfaatkan secara luas di berbagai sektor, termasuk bisnis, untuk memperoleh informasi berharga yang mendukung pengambilan keputusan. Namun, untuk menggali manfaat dari *big data*, diperlukan pendekatan analitis berupa arsitektur, algoritma, dan model analisis yang memadai (Fajriyah et al., 2022).

Salah satu pendekatan analitis tersebut adalah *data mining*, yaitu proses untuk menemukan sebuah pola tersembunyi agar mendapatkan pengetahuan baru di dalam sekumpulan data dilakukan penggalian data atau dikenal luas sebagai *data mining*. Konsep dasarnya adalah menentukan suatu informasi tersembunyi dalam sekumpulan data dengan melibatkan komputer dan manusia melalui proses manual ataupun otomatis yang bersifat berulang – ulang (Tarigan et al., 2022). *Data mining*

telah menjadi alat yang penting untuk mengubah sekumpulan data menjadi informasi. Alat ini dapat melakukan pekerjaan seperti memperkirakan, mengklasifikasikan hingga mengelompokkan data. Salah satu contoh penerapannya yaitu melakukan klasifikasi untuk mengelompokkan antara produk dan tidak laris (Nasrullah, 2021).

Di berbagai sektor, penerapan *data mining* telah menunjukkan hasil yang signifikan. Misalnya, dalam pengelolaan perpustakaan, teknik ini digunakan untuk menemukan pola kombinasi peminjaman buku yang sering dipinjam pada *database* peminjaman buku untuk mengumpulkan informasi yang kemudian berfungsi guna memberikan rekomendasi penempatan buku sehingga pengunjung perpustakaan mudah menemukannya. Hal ini juga membantu dalam pengambilan keputusan pengadaan buku agar sesuai dengan kebutuhan pengunjung (Andini et al., 2022). Penelitian lain juga menerapkan *data mining* dalam bidang kesehatan untuk mendeteksi penyakit secara dini, salah satunya untuk memprediksi terjadinya kanker paru (Sari et al., 2023). Masih di sektor kesehatan, implementasi lain yang relevan dengan penelitian ini adalah penerapan algoritma *Random Forest* dalam deteksi dini penyakit penyakit jantung yang tertuang dalam penelitian (Sumwiza et al., 2023) tentang pengembangan dan penyempurnaan model prediksi penyakit jantung (kardiovaskular) dengan menggunakan algoritma *Random Forest*.

Penyakit jantung merupakan salah satu penyebab utama kematian di dunia, termasuk di Indonesia. Menurut data dari Riskesdas tahun 2018, prevalensi penyakit jantung berdasarkan diagnosis dokter di Indonesia mencapai 1,5%, dengan angka tertinggi ditemukan di Provinsi Kalimantan Utara (2,2%), DI Yogyakarta (2%), dan Gorontalo (2%) (Kemenkes, 2021). Berdasarkan data yang dirilis oleh

(World Health Organization, 2021), penyakit jantung menyebabkan 17,8 juta kematian setiap tahunnya, atau sekitar satu dari tiga kematian di dunia disebabkan oleh penyakit jantung pada tahun 2019. Penyakit ini sering kali tidak menunjukkan gejala di awal, namun bisa tiba-tiba berujung pada serangan jantung mendadak yang fatal (Setyaji et al., 2018). Oleh karena itu, kebutuhan akan sistem pendukung pengambilan keputusan berbasis teknologi untuk membantu mengidentifikasi risiko penyakit jantung sedini mungkin menjadi sangat penting.

Salah satu tantangan terbesar dalam penanganan penyakit jantung adalah identifikasi risiko secara dini. Banyak pasien yang baru mengetahui bahwa mereka memiliki penyakit jantung setelah mengalami gejala berat atau serangan jantung pertama (Ndongala et al., 2022). Peran *data mining* menjadi krusial pada bagian analisis data dan parameter klinis, sistem dapat dilatih untuk mempelajari pola dari data pasien sebelumnya, dan kemudian digunakan untuk memprediksi kemungkinan risiko pada pasien baru (Premsmith & Ketmaneechairat, 2021).

Salah satu algoritma *data mining* yang bisa digunakan dalam klasifikasi *dataset* adalah *Random Forest*. Breiman dalam (Aprilia et al., 2021) mengusulkan pendekatan *Random Forest* sebagai pembelajaran mesin dengan banyak pohon keputusan. *Random Forest* merupakan metode klasifikasi dengan *ensemble learning* (Hidayat et al., 2023) yang membangun banyak pohon keputusan (*decision trees*) secara acak dari dataset (Depari et al., 2022), kemudian menggabungkan hasil prediksi dari masing-masing pohon untuk menghasilkan keputusan akhir yang lebih stabil dan akurat (Febtiawan et al., 2024).

Algoritma ini memiliki sejumlah keunggulan dibanding metode klasifikasi lain, di antaranya adalah kemampuannya dalam menangani dataset yang besar dengan fitur yang kompleks, kemampuannya dalam mengukur pentingnya tiap variabel prediktor, serta ketahanannya terhadap overfitting yang sering menjadi kendala dalam model prediktif (Ernawati et al., 2025). Random Forest juga mampu menangani data yang tidak seimbang (*imbalanced data*) dengan cukup baik (Mualfah et al., 2022), yang sering dijumpai dalam data medis, termasuk kasus penyakit jantung yang jumlah kasus positifnya lebih sedikit dibanding kasus negatif.

Dalam berbagai penelitian, Random Forest terbukti memberikan performa yang baik dalam prediksi diagnosis medis, termasuk untuk deteksi diabetes (Suryanegara et al., 2021), kanker (Chairunisa et al., 2020), dan penyakit kardiovaskular (Mualfah et al., 2022). Penelitian oleh (Firmansyah & Yulianto, 2023) menunjukkan bahwa penggunaan Random Forest dalam memprediksi risiko penyakit jantung memiliki tingkat akurasi yang tinggi.

Berdasarkan latar belakang tersebut, serta urgensi deteksi dini penyakit jantung yang semakin meningkat, dan potensi besar *data mining* dalam hal tersebut, maka penulis memutuskan untuk melakukan penelitian dengan judul “Penerapan *Data Mining* Untuk Identifikasi Risiko Terkena Penyakit Jantung Menggunakan *Random Forest*”.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian yang telah dijabarkan sebelumnya, peneliti menarik rumusan masalah yang akan diteliti, yaitu:

1. Bagaimana penerapan *random forest* untuk membuat model prediksi yang bisa mengidentifikasi tingkat risiko terkena penyakit jantung.
2. Bagaimana integrasi model prediksi untuk mengidentifikasi tingkat risiko penyakit jantung ke dalam sistem informasi berbasis situs web.

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk membangun Sistem Informasi berbasis situs web dan mengintegrasikannya dengan API yang memuat model prediksi risiko terkena penyakit jantung.

1.4 Manfaat Penelitian

Berikut manfaat yang diharapkan dari dilakukannya penelitian ini, yaitu:

1. Menambah wawasan dan pemahaman tentang penerapan *data mining* di sektor kesehatan.
2. Menyediakan sistem informasi yang dapat menjadi sarana edukasi sebagai upaya peningkatan kesadaran akan faktor risiko penyakit jantung.

1.5 Batasan Masalah

Guna menghindari pelebaran atau penyimpangan pokok masalah, penulis membatasi permasalahan yang akan dibahas. Batasan masalah yang diterapkan yaitu:

1. *Dataset* yang digunakan dalam penelitian ini adalah *heart disease dataset* yang diperoleh dari situs web Kaggle.
2. Atribut atau fitur yang digunakan dalam klasifikasi terbatas pada fitur – fitur yang tersedia dalam *dataset*, seperti kadar kolesterol, usia, jenis kelamin, detak jantung maksimum serta beberapa fitur lainnya.
3. Implementasi sistem dilakukan dalam bentuk situs web bersifat interaksi. Pengguna dapat memasukkan data dan memperoleh hasil prediksi risiko terkena penyakit jantung berdasarkan model prediksi yang telah dibangun.
4. Situs web yang dikembangkan tidak berbasis sistem informasi medis terintegrasi dan tidak digunakan untuk diagnosis medis, melainkan hanya sebagai sarana edukasi.
5. Penulis tidak membahas hal – hal yang terkait dengan keamanan sistem.

1.6 Sistematika Penulisan

Berikut ini sistematika penulisan dalam tugas akhir ini:

BAB I PENDAHULUAN

Bab ini menjelaskan latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, serta batasan penelitian.

BAB II TINJAUAN PUSTAKA

Bab ini berisi ulasan terhadap penelitian terdahulu yang relevan, teori – teori terkait *data mining* dan *random forest*, serta konsep dan teori lainnya.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan pendekatan penelitian yang dilakukan, termasuk metode pengumpulan data, teknik analisis data, algoritma yang digunakan, prosedur pengembangan sistem informasi berbasis situs web serta pengujian model.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil penelitian berupa hasil implementasi model prediksi risiko terkena penyakit jantung , evaluasi kinerja model dan hasil pengembangan sistem. Bab ini juga berisi pembahasan mengenai temuan penelitian dalam pembahasan.

BAB V KESIMPULAN DAN SARAN

Bab ini merupakan bab terakhir yang menyajikan kesimpulan dari keseluruhan penelitian termasuk capaian tujuan dari penelitian, serta memberikan saran untuk penelitian selanjutnya.

DAFTAR PUSTAKA

- Andini, Y., Hardinata, J. T., & Purba, Y. P. (2022). PENERAPAN DATA MINING TERHADAP TATA LETAK BUKU DI PERPUSTAKAAN SINTONG BINGEI PEMATANGSIANTAR MENGGUNAKAN METODE APRIORI. *Jurnal TIMES*, 11(1), 9–15. <http://ejournal.stmik-time.ac.id>
- Aprilia, W., Kurniawan, I., Baydhowi, M., & Haryati, T. (2021). Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest. *SISTEMASI: Jurnal Sistem Informasi*, 10(1), 163–171.
- Asadi, S., Roshan, S. E., & Kattan, M. W. (2021). Random forest swarm optimization-based for heart diseases diagnosis. *Journal of Biomedical Informatics*, 115, 1–13. <https://doi.org/10.1016/j.jbi.2021.103690>
- Asmara, J. (2019). Rancang Bangun Sistem Informasi Desa Berbasis Website (Studi Kasus Desa Netpala). *Jurnal Pendidikan Teknologi Informasi (JUKANTI)*, 2(1), 1–7.
- Awangga, R. M., & Khonsa, N. H. (2022). Analisis Performa Algoritma Random Forest dan Naive Bayes Multinomial pada Dataset Ulasan Obat dan Ulasan Film. *InComTech: Jurnal Telekomunikasi Dan Komputer*, 12(1), 60–70. <https://doi.org/10.22441/incomtech.v12i1.14770>
- Azzahra, F., Suarna, N., & Arie Wijaya, Y. (2024). Penerapan Algoritma Random Forest Dan Cross Validation Untuk Prediksi Data Stunting. *Kopertip : Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 8(1), 1–6. <https://doi.org/10.32485/kopertip.v8i1.238>

- Bachtiar, L., Gustaman, R. A., & Maywati, S. (2023). FAKTOR RISIKO YANG BERHUBUNGAN DENGAN KEJADIAN PENYAKIT JANTUNG KORONER (PJK) (Analisis Data Sekunder di Rumah Sakit Umum Daerah Kabupaten Subang). *Jurnal Kesehatan Komunitas Indonesia*, 19(1), 52–60.
- Binsawad, M. (2024). Enhancing kidney disease prediction with optimized forest and ECG signals data. *Heliyon*, 10, 1–11. <https://doi.org/10.1016/j.heliyon.2024.e30792>
- Casonatto, R. A., Souza, T. D. P. G., & Mariano, A. M. (2024). Quality and Risk Management in Data Mining: A CRISP-DM Perspective. *Procedia Computer Science*, 242, 161–168. <https://doi.org/10.1016/j.procs.2024.08.257>
- Chairunisa, R., Adiwijaya, & Astuti, W. (2020). Perbandingan CART dan Random Forest untuk Deteksi Kanker berbasis Klasifikasi Data Microarray. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(5), 805–812.
- Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. A. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, 2(November 2021), 100016. <https://doi.org/10.1016/j.health.2022.100016>
- Depari, D. H., Widiastiwi, Y., & Santoni, M. M. (2022). Perbandingan Model Decision Tree , Naive Bayes dan Random Forest untuk Prediksi Klasifikasi Penyakit Jantung. *Jurnal Informatik*, 18(3), 239–248.
- Ernawati, A., Khairul, Sitorus, Z., Iqbal, M., & Nasution, D. (2025). Penerapan Data Mining Untuk Klasifikasi Penduduk Miskin Di Kabupaten Labuhanbatu

- Menggunakan Random Forest Dan. *Bulletin of Information Technology (BIT)*, 6(2), 23–35. <https://doi.org/10.47065/bit.v5i2.1783>
- Eska, J. (2016). Penerapan Data Mining Untuk Prediksi Penjualan Walpaper Menggunakan Algoritma C4.5. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 2(2), 9–13.
- Fajriyah, N., Setiawan, W., Dewi, E., & Duha, T. (2022). IMPLEMENTASI TEKNOLOGI BIG DATA DI ERA DIGITAL. *Jurnal Informatika*, 1(1), 1–7.
- Febtiawan, E. P., Akbar, L. A. S. I., & Rachman, A. S. (2024). Forecasting Produksi Energi Photovoltaic Menggunakan Algoritma Random Forest Classification. *Journal of Information System Research*, 5(4), 1053–1062. <https://doi.org/10.47065/josh.v5i4.5514>
- Firmansyah, & Yulianto, A. (2023). Prediksi Penyakit Jantung Menggunakan Algoritma Random Forest. *Jurnal Minfo Polgan*, 12(2), 2239–2246.
- Ghosh, A., & Jana, S. (2022). A Study on Heart Disease Prediction using Different Classification Models based on Cross Validation Method. *International Journal of Engineering Research & Technology (IJERT)*, 11(06), 67–71. www.ijert.org
- Hidayat, Sunyoto, A., & Fatta, H. Al. (2023). Klasifikasi Penyakit Jantung Menggunakan Random Forest Clasifier. *Jurnal Sistem Komputer Dan Kecerdasan Buatan*, VII(1), 31–40.
- Huber, S., Wiemer, H., Schneider, D., & Ihlenfeldt, S. (2019). DMME: Data mining methodology for engineering applications - A holistic extension to the CRISP-

DM model. *Procedia CIRP*, 79, 403–408.

<https://doi.org/10.1016/j.procir.2019.02.106>

Kemenkes. (2021). *Penyakit Jantung Koroner Didominasi Masyarakat Kota*.

Kurniawan, H. (2020). Deteksi Twitter Bot Menggunakan Klasifikasi Decision Tree. *Jurnal Sustainable : Jurnal Hasil Penelitian Dan Industri Terapan*, 9(1), 31–37.

Marisa, F., Manukar, A. L., & Akhriza, T. M. (2021). *Data Mining Konsep dan Penerapannya*. CV Budi Utama.

Mualfah, D., Fadila, W., & Firdaus, R. (2022). Teknik SMOTE untuk mengatasi imbalance data pada deteksi penyakit stroke menggunakan algoritma random forest. *Jurnal Computer Science and Information Technology (CoSciTech)*, 3(2), 107–113.

Mujiani, S., & Mardhiyah, K. (2019). Perancangan dan Penerapan Sistem Informasi Akuntansi Berbasis Website Pada Fakultas Ekonomi dan Bisnis Universitas Islam As-Syafi 'iyah. *Jurnal Akuntansi Dan Keuangan*, 1(1), 34–47.

Naomi, W. S., Picauly, I., & Toy, S. M. (2021). Faktor Risiko Kejadian Penyakit Jantung Koroner (Studi Kasus di RSUD Prof. Dr. W. Z. Johannes Kupang). *Media Kesehatan Masyarakat*, 3(1), 99–107.
<https://doi.org/10.35508/mkm.v3i1.3622>

Nasrullah, A. H. (2021). IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI PRODUK LARIS. *Jurnal Ilmiah Ilmu Komputer*, 7(2). <http://ejournal.fikom-unasman.ac.id>

- Ndongala, N. J., Maepa, C., Nyondo, E., Amstutz, A., & du Reau de la Gaignonnière, B. (2022). Etiology, Characteristics and Occurrence of Heart Diseases in Rural Lesotho (ECHOLesotho): A Retrospective Echocardiography Cohort Study. *PLOS ONE*, 17(12), 1–11. <https://doi.org/10.1371/journal.pone.0278406>
- Nugraha, A. F., Aziza, R. F. A., & Pristyanto, Y. (2022). Penerapan metode Stacking dan Random Forest untuk Meningkatkan Kinerja Klasifikasi pada Proses Deteksi Web Phishing. *Jurnal Infomedia: Teknik Informatika, Multimedia & Jaringan*, 7(1), 39–44.
- Nugroho, A. (2022). Analisa Splitting Criteria Pada Decision Tree dan Random Forest untuk Klasifikasi Evaluasi Kendaraan. *JSITIK*, 1(1), 41–49.
- Nuraeni, N. (2024). Klasifikasi Data Mining untuk Prediksi Penyakit Kardiovaskular. *Jurnal Teknik Informasi Dan Komputer*, 7(1), 161–170. <https://doi.org/10.37600/tekinkom.v7i1>.
- Nurchahyo, J. A., & Sasongko, T. B. (2023). Hyperparameter Tuning Algoritma Supervised Learning untuk Klasifikasi Keluarga Penerima Bantuan Pangan Beras. *Indonesian Journal of Computer Science*, 12(3), 1351–1365. <https://doi.org/10.33022/ijcs.v12i3.3254>
- Plotnikova, V., Dumas, M., & Milani, F. P. (2022). Applying the CRISP-DM data mining process in the financial services industry: Elicitation of adaptation requirements. *Data and Knowledge Engineering*, 139, 1–17. <https://doi.org/10.1016/j.datak.2022.102013>

- Premsmith, J., & Ketmaneechairat, H. (2021). A Predictive Model for Heart Disease Detection Using Data Mining Techniques. *Journal of Advances in Information Technology*, 12(1), 14–20. <https://doi.org/10.12720/jait.12.1.14-20>
- Putri, N. B., & Wijayanto, A. W. (2022). Analisis Komparasi Algoritma Klasifikasi Data Mining Dalam Klasifikasi Website Phishing. *Komputika : Jurnal Sistem Komputer*, 11(1), 59–66. <https://doi.org/10.34010/komputika.v11i1.4350>
- Rabbani, D. A., & Najicha, F. U. (2023). Pengaruh Perkembangan Teknologi terhadap Kehidupan dan Interaksi Sosial Masyarakat Indonesia. *Research Gate*, 10(3), 1–11. <https://www.researchgate.net/publication/375525102>
- Rachmawati, C., Martini, S., & Artanti, K. D. (2021). Analisis Faktor Risiko Modifikasi Penyakit Jantung Koroner Di Rsu Haji Surabaya Tahun 2019. *Media Gizi Kesmas*, 10(1), 47–55. <https://doi.org/10.20473/mgk.v10i1.2021.47-55>
- Rahman, R., & Sutanto, F. A. (2023). Data Mining Untuk Memprediksi Tingkat Kepuasan Konsumen Gojek Menggunakan Algoritma Naive Bayes. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 18(1), 8–18. <https://doi.org/10.35969/interkom.v18i1.280>
- Ramadhon, R. N., Ogi, A., Agung, A. P., Putra, R., Febrihartina, S. S., & Firdaus, U. (2024). Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank. *Karimah Tauhid*, 3(2), 1860–1874. <https://doi.org/10.30997/karimahtauhid.v3i2.11952>
- Riani, A., Susianto, Y., & Rahman, N. (2019). Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Menggunakan Metode Naive Bayes. *Journal of*

Innovation Information Technology and Application (JINITA), 1(01), 25–34.
<https://doi.org/10.35970/jinita.v1i01.64>

Riski, A. (2019). Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Penderita Penyakit Jantung. *Jurnal Teknik Informatika Kaputama (JTIK)*, 3(1), 22–28.

Rohimah, S., Rosdiana, N., & Ayuningsih, R. (2024). Pelatihan Kader Usaha Kesehatan Sekolah dalam Deteksi Dini Faktor Risiko Penyakit Tidak Menular : Pencegahan Kegawatdaruratan Penyakit Kardiovaskular. *JPKMU: Jurnal Pengabdian Kepada Masyarakat Kesehatan Galuh*, 1(2), 111–120.

Saefudin, M., Megawaty, D. A., Alita, D., Arundaa, R., & Tenda, E. (2023). Penerapan Framework Laravel Pada Sistem Informasi Posyandu Berbasis Website. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 4(2), 213–220.
<https://doi.org/10.33365/jatika.v4i2.2600>

Santoso, H., Putri, R. A., & Sahbandi, S. (2023). Deteksi Komentar Cyberbullying pada Media Sosial Instagram Menggunakan Algoritma Random Forest. *Jurnal Manajemen Informatika (JAMIKA)*, 13(1), 62–72.
<https://doi.org/10.34010/jamika.v13i1.9303>

Santoso, R. E., Prawiyogi, A. G., Rahardja, U., Oganda, F. P., & Khofifah, N. (2022). Penggunaan dan Manfaat Big Data dalam Konten Digital. *ADI Bisnis Digital Interdisiplin*, 3(2), 157–160.
<https://doi.org/https://doi.org/10.34306/abdi.v3i2.836>

- Saraswati, D. (2024). Inovasi Pelayanan Kesehatan: Deteksi Dini Penyakit Jantung Koroner melalui Posbindu PTM. *Jurnal Kesehatan Dan Kebidanan Nusantara*, 2(1), 10–16. <https://doi.org/10.69688/jkn.v2i1.81>
- Saraswati, D., & Lina, N. (2020). Faktor Risiko Penyakit Jantung Pada Masyarakat Di Pos Pembinaan Terpadu (Posbindu) Puskesmas Cibeureum. *Journal Health & Science : Gorontalo Journal Health and Science Community*, 4(1), 1–9. <https://doi.org/10.35971/gojhes.v2i1.4426>
- Sari, L., Romadloni, A., & Listyaningrum, R. (2023). Penerapan Data Mining dalam Analisis Prediksi Kanker Paru Menggunakan Algoritma Random Forest. *Infotekmesin*, 14(1), 155–162. <https://doi.org/10.35970/infotekmesin.v14i1.1751>
- Sawu, S. D., Prayitno, A. A., & Wibowo, Y. I. (2022). Analisis Faktor Risiko pada Kejadian Masuk Rumah Sakit Penyakit Jantung Koroner di Rumah Sakit Husada Utama Surabaya. *Jurnal Sains Dan Kesehatan*, 4(1), 10–18. <https://doi.org/10.25026/jsk.v4i1.856>
- Scandea, E. I., Sugiarto, M. A. R., Lestari, F., & Hartanti, D. (2023). Penerapan Data Mining Untuk Menganalisis Data Faktor Resiko Penyakit Jantung Menggunakan Metode Logistic Regression. *Prosiding Seminar Nasional Teknologi Informasi Dan Bisnis (SENATIB)*, 683–688.
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181, 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>

- Setyaji, D. Y., Prabandari, Y. S., & Gunawan, I. M. A. (2018). Aktivitas fisik dengan penyakit jantung koroner di Indonesia. *Jurnal Gizi Klinik Indonesia*, 14(3), 115–121.
- Sitanggang, D., Nicholas, N., Wilson, V., Sinaga, A. R. A., & Simanjuntak, A. D. (2022). Implementasi Data Mining Untuk Memprediksi Penyakit Jantung Menggunakan Metode K-Nearest Neighbor Dan Logistic Regression. *Jurnal Teknik Informasi Dan Komputer (Tekinkom)*, 5(2), 493–501. <https://doi.org/10.37600/tekinkom.v5i2.698>
- Sulianta, F. (2023). *Buku Dasar Data Mining from A to Z*.
- Sumwiza, K., Twizere, C., Rushingabigwi, G., Bakunzibake, P., & Bamurigire, P. (2023). Enhanced cardiovascular disease prediction model using random forest algorithm. *Informatics in Medicine Unlocked*, 41, 1–9. <https://doi.org/10.1016/j.imu.2023.101316>
- Suryanegara, G. A. B., Adiwijaya, & Purbolaksono, M. D. (2021). Peningkatan Hasil Klasifikasi pada Algoritma Random Forest untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 114–122.
- Susanti, N., Zahara, A., Darus, nanda F., & Zulaila. (2024). Faktor Risiko Yang Berhubungan Dengan Kejadian Penyakit Jantung Koroner : Literatur Riview. *Jurnal Kesehatan Tambusai*, 5(2), 4217–4222. <https://doi.org/10.37058/jkki.v19i1.6862>

- Susanti, Z., Sirait, P., & Panjaitan, E. S. (2023). Peningkatan Kinerja Random Forest Melalui Seleksi Fitur Secara Pca Untuk Mendeteksi Penyakit Diabetes Tahap Awal. *Jurnal Sains Dan Teknologi*, 4(3), 51–56.
- Susilo, H., Abdillah, N., Ikhsan, M., & Diana Morika, H. (2023). Analisa Dan Perancangan Sistem Informasi Booking Antrian Pelayan Pada Klinik Medika Saintika Berbasis Website. *Jurnal Kesehatan Medika Saintika*, 14(1), 344–352.
- Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., M.Safii, & Winanjaya, R. (2022). IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA APRIORI DALAM MENENTUKAN PERSEDIAAN BARANG (STUDI KASUS : TOKO SINAR HARAHAP). *Jurnal Sistem Informasi, Teknologi Informasi, Dan Komputer*, 12(2), 51–61.
<https://jurnal.umj.ac.id/index.php/just-it/index>
- Wahyudin, Y., & Rahayu, D. N. (2020). Analisis Metode Pengembangan Sistem Informasi Berbasis Website: A Literatur Review. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 15(3), 119–133. <https://doi.org/10.35969/interkom.v15i3.74>
- Widyanti, D., & Widyastuti, C. S. (2020). Faktor Risiko Terjadinya Penyakit Jantung Koroner pada Pasien di Unit Rawat Jalan Rumah Sakit Panti Rapih Yogyakarta. *Jurnal Keperawatan I Care*, 1(1), 26–42.
<https://doi.org/10.46668/jurkes.v1i1.34>
- World Health Organization. (2021). *Cardiovascular diseases (CVDs)*.

Zahra, R. S., Virgantara Putra, O., & Putri, R. U. (2022). Rancang Bangun Sistem Informasi Telemedicine Pada Darussalam Medical Center (DMC) Menggunakan Framework Laravel Design and Build a Telemedicine Information System at Darussalam Medical Center (DMC) Using the Laravel Framework. *Seminar Nasional Hasil Penelitian & Pengabdian Masyarakat Bidang Ilmu Komputer*, 84–91.