

**PENERAPAN ALGORITMA RELIEFF
DAN *BACKPROPAGATION NEURAL NETWORK* (BNN)
DALAM PENANGANAN DATA HILANG
DAN KLASIFIKASI PENYAKIT JANTUNG**

SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Sains Bidang Studi Matematika**

Oleh :

Henny Nur Syafitri

NIM 08011281823045



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA**

2022

LEMBAR PENGESAHAN

**PENERAPAN ALGORITMA RELIEFF
DAN BACKPROPAGATION NEURAL NETWORK (BNN)
DALAM PENANGANAN DATA HILANG DAN KLASIFIKASI PENYAKIT JANTUNG**

SKRIPSI

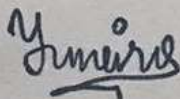
**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Sains Bidang Studi Matematika**

Oleh

HENNY NUR SYAFITRI

NIM. 08011281823045

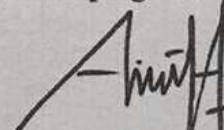
Pendamping Kedua



Irmeilvana, M.Si

NIP. 19740517 1999032003

**Indralaya, Mei 2022
Pendamping Utama**



Anita Desiani, S.S., M.Kom

NIP. 19771211 2003122002



PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Henny Nur Syafitri

NIM : 08011281823045

Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan Alam/Matematika

Menyatakan bahwa skripsi ini adalah hasil karya saya sendiri dan karya ilmiah ini belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar kesarjanaan srata (S1) dari Universitas Sriwijaya maupun perguruan tinggi lain. Semua informasi yang dimuat dalam skripsi ini berasal dari penullis lain baik yang dipublikasikan atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, 30 Juni 2022

Penulis



Henny Nur Syafitri

NIM.08011281823045

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai Civitas Akademik Universitas Sriwijaya, yang bertanda tangan di bawah ini :

Nama Mahasiswa : Henny Nur Syafitri
NIM : 08011281823045
Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan Alam/Matematika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikann kepada Universitas Sriwijaya “Penerapan Algoritma ReliefF dan *Backpropagation Neural Network* (BNN) dalam Penanganan Data Hilang dan Klasifikasi Penyakit Jantung”. Dengan hak bebas royalty non-eksklusif ini Universitas Sriwijaya berhak menyimpan, mengalih, edit/memformatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir atau skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Indralaya, 30 Juni 2022

Penulis



Henny Nur Syafitri
NIM.08011281823045

LEMBAR PERSEMBAHAN

Kupersembahkan skripsi ini untuk:

Yang Maha Kuasa Allah Subhanahu Wa Ta'ala,

Kedua orang tuaku tersayang,

Adik Laki-lakiku,

Keluarga besarku,

Semua guru dan dosenku,

Sahabat-sahabatku,

Almamaterku

Moto

“Berusahalah semaksimal mungkin, lalu serahkan hasilnya pada Allah SWT”

KATA PENGANTAR

Puji syukur kehadiran Allah Subhanahu wa Ta'ala yang telah memberikan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Algoritma Relieff Dan *Backpropagation Neural Network* (BNN) dalam Penanganan Data Hilang dan Klasifikasi Penyakit Jantung” sebagai salah satu syarat memperoleh gelar sarjana sains bidang studi Matematika di Fakultas MIPA Universitas Sriwijaya.

Penulis menyadari bahwa proses pembuatan skripsi ini merupakan proses pembelajaran yang sangat berharga serta tak lepas dari kekurangan dan keterbatasan. Dengan segala hormat dan kerendahan hati, penulis mengucapkan terima kasih dan penghargaan kepada:

1. Kedua orang tuaku tercinta, Bapak **Hardi** dan Ibu **Endang Lestari**, yang tak pernah lelah merawat, mendidik, menuntun, memberi nasehat, semangat serta doa untuk penulis. Terima kasih atas segala perjuangan, pengorbanan, serta kasih sayang hingga detik ini dan sampai kapanpun.
2. Bapak **Drs. Sugandi Yahdin, M.M** selaku Ketua Jurusan Matematika FMIPA Universitas Sriwijaya yang telah memberikan arahan dan motivasi kepada penulis selama proses perkuliahan.
3. Ibu **Dr. Dian Cahyawati Sukanda, M.Si** selaku Sekretaris Jurusan Matematika FMIPA Universitas Sriwijaya yang telah mengarahkan urusan akademik kepada penulis.

4. Ibu **Anita Desiani, S.Si., M.Kom** selaku dosen pembimbing utama yang telah bersedia memberikan waktu, tenaga, pikiran, nasehat, dan motivasi untuk memberikan bimbingan dan pengarahan selama proses pembuatan skripsi, kompetisi, dan perjalanan perkuliahan ini.
5. Ibu **Irmeilyana, M.Si** selaku dosen pembimbing pendamping sekaligus dosen pembimbing akademik yang telah bersedia memberikan waktu, tenaga, pikiran, nasehat, dan motivasi untuk memberikan bimbingan dan pengarahan selama proses pembuatan skripsi serta proses perkuliahan dengan penuh pengertian dan kesabaran.
6. Ibu **Dr. Ir. Herlina Hanum, M.Si** dan Ibu **Eka Susanti, M.Sc** selaku dosen pembahas dan penguji yang telah memberikan tanggapan, kritik, dan saran yang sangat bermanfaat untuk perbaikan dan penyelesaian skripsi ini.
7. **Seluruh Dosen di Jurusan Matematika FMIPA** yang telah memberikan ilmu, nasihat, motivasi, serta bimbingan selama proses perkuliahan.
8. Pak **Irwansyah** selaku admin dan Ibu **Hamidah** selaku pegawai tata usaha Jurusan Matematika FMIPA yang telah membantu penulis selama perkuliahan.
9. **Seluruh guru** yang telah memberikan ilmu yang bermanfaat hingga mengantarkan penulis pada pendidikan ini.
10. Adik laki-lakiku tersayang, **Hendy Halim Prasetyo**, yang selalu mendoakan dan memberikan perhatian kepada penulis, beserta keluarga besar yang selalu mendukung penulis.

11. **Semua sahabat seperjuangan** dalam masa perkuliahan dan proses skripsi.
Terima kasih sudah menjadi orang-orang baik di sekeliling penulis yang selalu mendukung, membantu dengan tulus, dan memberi energi positif.
12. **Keluarga Matematika 2018, Tim PHP2D, BPH Himastik Beraksi dan Gelora Karya**, dan **rekan-rekan perlombaan** selama perkuliahan.
13. Kakak-kakak tingkat **angkatan 2016** dan **2017** serta adik-adik tingkat **angkatan 2019** dan **2020**, terima kasih atas segala kebaikan dan bantuan.
14. Semua pihak yang tidak dapat penulis sebutkan satu per satu. Semoga segala kebaikan yang diberikan mendapatkan balasan terbaik dari Allah SWT.

Semoga skripsi ini dapat menambah pengetahuan dan bermanfaat bagi mahasiswa/mahasiswi Jurusan Matematika Fakultas dan Ilmu Pengetahuan Alam Universitas Sriwijaya dan semua pihak yang memerlukan.

Indralaya, Mei 2022

Penulis

**APPLICATION OF RELIEFF ALGORITHM
AND BACKPROPAGATION NEURAL NETWORK (BNN)
IN HANDLING OF MISSING VALUE
AND CLASSIFICATION OF HEART DISEASE**

By:

**Henny Nur Syafitri
08011281823045**

ABSTRACT

Classification is the process of classifying an object based on a certain model. This process is often used in various fields, such as health, environment, and data research. One application of classification in the health sector is to predict factors for heart disease. The dataset used to predict cardiovascular disease factors comes from the University of California Irvine (UCI). This dataset has a weakness there is missing data in some features. Missing data can be overcome by applying the ReliefF algorithm, where each feature is sorted from the largest weight to the smallest weight. The advantage of the ReliefF algorithm is that it does not limit the data types used and can effectively deal with multiclass problems, missing data, and noise tolerance. In this research, the application of the ReliefF algorithm and testing using the Backpropagation Neural Network (BNN) method was carried out to determine the best features of the predictive factors for heart disease. Before applying the ReliefF algorithm in the detection of heart disease, the accuracy, precision, and recall values obtained were 77.21%, 72.31%, and 76.60%. Meanwhile, after applying the ReliefF algorithm, the performance evaluation results obtained were 84.01%, 80.09%, and 84.76%. Based on the performance results obtained through testing using the Backpropagation Neural Network (BNN) method, these values indicate that the ReliefF algorithm can be used to select the most influential and good features to increase the level of accuracy in predicting heart disease patients.

Keywords : Classification, heart disease, missing value, ReliefF algorithm, Backpropagation Neural Network (BNN)

**PENERAPAN ALGORITMA RELIEFF
DAN BACKPROPAGATION NEURAL NETWORK (BNN)
DALAM PENANGANAN DATA HILANG
DAN KLASIFIKASI PENYAKIT JANTUNG**

Oleh:

**Henny Nur Syafitri
08011281823045**

ABSTRAK

Klasifikasi adalah proses penggolongan suatu objek berdasarkan model tertentu. Proses ini sering digunakan dalam berbagai bidang, seperti bidang kesehatan, lingkungan, dan penyelidikan data. Salah satu pengaplikasian klasifikasi dalam bidang kesehatan ialah untuk memprediksi faktor penyakit jantung. *Dataset* yang digunakan untuk memprediksi faktor penyakit jantung berasal dari *University of California Irvine* (UCI). *Dataset* ini mempunyai kelemahan yaitu terdapat data hilang di beberapa fitur. Data hilang dapat diatasi dengan menerapkan algoritma ReliefF, dimana masing-masing fitur diurutkan dari bobot yang terbesar sampai bobot yang terkecil. Kelebihan dari algoritma ReliefF ialah tidak membatasi tipe data yang digunakan serta dapat secara efektif menangani masalah multikelas, data hilang, dan toleransi *noise*. Pada penelitian ini, dilakukan penerapan algoritma ReliefF dan pengujian menggunakan metode Backpropagation Neural Network (BNN) guna mengetahui fitur-fitur terbaik dari faktor-faktor prediksi penyakit jantung. Sebelum penerapan algoritma ReliefF dalam deteksi penyakit jantung, nilai *accuracy*, *precision*, dan *recall* yang didapat sebesar 77,21%, 72,31%, dan 76,60%. Sedangkan setelah diterapkan algoritma ReliefF, hasil evaluasi kinerja yang di dapat sebesar 84,01%, 80,09%, dan 84,76%. Berdasarkan hasil kinerja yang didapat melalui pengujian menggunakan metode *Backpropagation Neural Network* (BNN), nilai-nilai tersebut menunjukkan bahwa algoritma ReliefF dapat digunakan untuk memilih fitur yang paling berpengaruh dan baik untuk menaikkan tingkat *accuracy* pada prediksi pasien penyakit jantung.

Kata Kunci : Klasifikasi, penyakit jantung, data hilang, algoritma ReliefF, *Backpropagation Neural Network* (BNN)

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN PENGESAHAN	ii
LEMBAR PERSEMBAHAN	iii
KATA PENGANTAR	vi
ABSTRACT	ix
ABSTRAK	x
DAFTAR ISI	xi
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	4
1.3 Pembatasan Masalah.....	4
1.4 Tujuan.....	4
1.5 Manfaat.....	5
BAB II TINJAUAN PUSTAKA	6
2.1 Data Hilang.....	6
2.2 <i>Feature Selection</i>	6
2.3 <i>Manhattan Distance</i>	7
2.4 Algoritma ReliefF.....	7
2.5 Klasifikasi.....	10
2.6 <i>Backpropagation</i>	10
2.7 <i>Confusion Matrix</i>	12
BAB III METODOLOGI PENELITIAN	15
3.1 Tempat.....	15
3.2 Waktu.....	15
3.3 Alat.....	15
3.4 Metode Penelitian.....	15
BAB IV HASIL DAN PEMBAHASAN	20
4.1 Deskripsi Data.....	20
4.2 Model Perhitungan Manual.....	22
4.2.1 Perhitungan Manual dengan Algoritma ReliefF.....	22
4.2.2 Perhitungan Manual dengan Metode <i>Backpropagation</i>	30
4.3 Hasil <i>Feature Selection</i> Algoritma ReliefF.....	37
4.4 Klasifikasi <i>Backpropagation</i> pada Hasil <i>Feature Selection</i>	38
4.5 Klasifikasi pada Data Sebelum Penerapan <i>Feature Selection</i>	39
4.6 Klasifikasi pada Data Setelah Penerapan <i>Feature Selection</i>	43
4.7 Analisis Hasil.....	46
BAB V KESIMPULAN DAN SARAN	51
5.1 Kesimpulan.....	51
5.2 Saran.....	51
DAFTAR PUSTAKA	53

DAFTAR TABEL

Tabel 2.1 <i>Confusion matrix</i>	13
Tabel 2.2 Nilai evaluasi kinerja model	14
Tabel 4.1 <i>Dataset</i> pasien penyakit jantung	20
Tabel 4.2 Keterangan fitur	20
Tabel 4.3 Data contoh perhitungan manual menggunakan algoritma ReliefF	22
Tabel 4.4 Nilai bobot “ <i>cp</i> ” pada tiap iterasi dengan $m = 10$ dan $k = 2$	25
Tabel 4.5 Nilai bobot “ <i>restecg</i> ” pada tiap iterasi dengan $m = 10$ dan $k = 2$	27
Tabel 4. 6 Nilai bobot “ <i>oldpeak</i> ” pada tiap iterasi dengan $m = 10$ dan $k = 2$	29
Tabel 4.7 Hasil bobot untuk setiap fitur.....	29
Tabel 4.8 Bobot masukan pelatihan contoh perhitungan manual dari <i>input layer</i> menuju <i>hidden layer</i>	30
Tabel 4.9 Bobot masukan pelatihan contoh perhitungan manual dari <i>hidden layer</i> menuju <i>output layer</i>	30
Tabel 4.10 Tabel bobot terbaru pada <i>output layer</i> dan <i>hidden layer</i>	36
Tabel 4.11 Nilai bobot setiap fitur dengan algoritma ReliefF	37
Tabel 4.12 Hasil perbandingan nilai <i>accuracy</i> pada setiap seleksi fitur.....	38
Tabel 4.13 <i>Confusion matrix</i> metode <i>backpropagation</i> pada data sebelum <i>feature selection</i>	39
Tabel 4.14 Hasil evaluasi data dengan metode <i>backpropagation</i> sebelum <i>feature selection</i>	43
Tabel 4.15 <i>Confusion matrix</i> metode <i>backpropagation</i> pada data setelah <i>feature selection</i>	43
Tabel 4.16 Hasil evaluasi data dengan metode <i>backpropagation</i> setelah <i>feature selection</i>	46
Tabel 4.17 Perbandingan hasil <i>feature selection</i> algoritma ReliefF pada data.....	46
Tabel 4.18 Hasil perhitungan rata-rata nilai <i>precision</i> dan nilai <i>recall</i> pada label sehat dan label sakit	48
Tabel 4. 19 Perbandingan hasil penelitian dengan penelitian lain.....	49

DAFTAR GAMBAR

Gambar 2.1 Arsitektur metode <i>backpropagation</i>	10
Gambar 3.1 <i>Flowchart</i> proses <i>backpropagation</i>	19

BAB I

PENDAHULUAN

1.1 Latar Belakang

Klasifikasi adalah proses penggolongan suatu objek berdasarkan model tertentu (Deng *et al.*, 2016). Klasifikasi sering digunakan dalam berbagai bidang, seperti bidang kesehatan, lingkungan, dan penyelidikan data (Mohan *et al.*, 2019). Salah satu pengaplikasian klasifikasi dalam bidang kesehatan ialah untuk memprediksi penyakit jantung. Penyakit jantung merupakan salah satu penyebab kematian terbesar di dunia saat ini (Mohan *et al.*, 2019). Penyakit ini disebabkan oleh penyempitan pembuluh darah arteri yang berfungsi menyalurkan nutrisi dan oksigen ke jantung (Johnson *et al.*, 2019). Beberapa faktor yang mempengaruhi penyakit jantung ialah tingkat kolestrol tinggi, tekanan darah tinggi, obesitas, riwayat keturunan, dan lain-lain (Yang & Garibaldi, 2015). Hal-hal yang dapat dilakukan guna mencegah terjadinya penyakit jantung antara lain mengubah gaya hidup menjadi lebih sehat, melakukan terapi antiplatelet, menurunkan berat badan, dan melakukan kontrol glikemik (Oktay *et al.*, 2018). Penyakit jantung dapat diprediksi dengan melakukan tes elektrokardiogram, guna memprediksi gejala awal dari penyakit jantung dengan cara mengukur dan merekam aktivitas jantung (Otto *et al.*, 2021).

Salah satu *dataset* yang dapat digunakan untuk memprediksi faktor penyakit jantung yaitu *dataset* yang berasal dari *University of California Irvine (UCI) Machine Learning Repository*. *Dataset* tersebut dapat diakses melalui laman (<https://archive.ics.uci.edu/ml/datasets/heart+disease>) yang berisi data dari

Cleveland, Switzerland, VA Long Beach dan Hungarian yang terdiri atas 76 fitur, namun hanya 14 fitur yang dipublikasikan secara umum (Shah *et al.*, 2017). Pada *dataset* penyakit jantung, klasifikasi terbagi dua yaitu menderita penyakit jantung (sakit) dan tidak menderita penyakit jantung (sehat). *Dataset* tersebut mengalami data hilang, dimana pada beberapa fitur tidak terdapat nilai yang lengkap, sehingga mengakibatkan banyak informasi yang hilang (Desiani *et al.*, 2021).

Data hilang dapat terjadi karena berbagai alasan, seperti kurang tepat saat dilakukan pengukuran dan kurang akurat saat pengisian kuisioner, sehingga mempengaruhi hasil kinerja dalam mengenali pola pada masalah klasifikasi (Pedersen *et al.*, 2017). Salah satu cara yang diterapkan untuk mengatasi data hilang ialah penghapusan (*cleansing*), yaitu dengan cara menghapus baris yang tidak lengkap dan menghapus fitur yang tidak penting (Lang & Little, 2018). Metode yang dapat digunakan untuk memilih fitur-fitur yang paling penting ialah algoritma ReliefF. Algoritma ReliefF merupakan salah satu model pemilihan fitur berbasis *filter* yang diterapkan secara luas dan memiliki efisiensi klasifikasi yang besar (Sun *et al.*, 2019). Kelebihan dari algoritma ini adalah tidak membatasi tipe data yang digunakan serta dapat secara efektif menangani masalah multikelas, data hilang, dan toleransi *noise* (Osanaiye *et al.*, 2019). Beberapa penelitian lain menggunakan algoritma ReliefF menunjukkan hasil yang sangat baik, seperti penelitian yang dilakukan oleh Baliarsingh *et al.* (2019) mengenai data penyakit Colon tumor. Hasil *accuracy* yang diperoleh sebesar 84,26%. Penelitian lain dilakukan oleh Q. Liu *et al.* (2017) dengan menggabungkan algoritma ReliefF dan

Random Forest (RF) untuk memperbaiki diagnosa dalam bidang kesehatan menghasilkan *accuracy* sebesar 91,82%.

Pada klasifikasi, pemilihan metode juga menentukan kinerja klasifikasi, salah satu metode yang saat ini populer digunakan ialah metode *Artificial Neural Network* (ANN) (Quaglio *et al.*, 2020). Metode ini mempunyai beberapa kelebihan yaitu memiliki kemampuan *self-learning* dan *self-adaptation*, serta memiliki kemampuan untuk memetakan hubungan antara *input* dan *output* dari suatu sistem secara non-linear (Jawad *et al.*, 2021). Salah satu metode dari *Artificial Neural Network* (ANN) adalah metode *backpropagation* (Yamashita *et al.*, 2018). Tujuan *training* menggunakan *backpropagation* ialah untuk menghasilkan bobot baru dari masing-masing fitur dan meminimalisir jumlah error antara nilai aktual dan nilai prediksi (Mohd Nawi *et al.*, 2021). Kelebihan dari *backpropagation* yaitu memiliki sifat komputasi yang baik terutama saat mengolah data dengan skala yang besar dan kompleks (Setti & Wanto, 2019).

Zhang *et al.* (2016) melakukan penelitian menggunakan *Backpropagation Neural Network* (BNN) tentang penyakit jantung dan kanker kolorektal, dengan hasil *accuracy* sebesar 75%, *precision* sebesar 70% dan *recall* sebesar 80%. Penelitian lain dilakukan oleh Mhatre & Varma (2019) yang memprediksi penyakit jantung menggunakan *backpropagation*. Penelitian ini menghasilkan nilai *accuracy* sebesar 78,76% serta tanpa menghitung ukuran kinerja evaluasi lainnya. Selain itu, Al-Barzinji *et al.* (2020) menggunakan metode *backpropagation* untuk memprediksi penyakit jantung yang menghasilkan nilai *accuracy* sebesar 82,17%, *precision* sebesar 81%, dan *recall* sebesar 79%.

Berdasarkan kelebihan dari algoritma ReliefF dan metode *backpropagation* yang sudah dijelaskan sebelumnya, maka penelitian ini menentukan fitur-fitur yang terpenting dari *dataset* UCI menggunakan algoritma ReliefF, kemudian mengukur kinerja klasifikasi dengan menggunakan metode *backpropagation* dengan mengukur *accuracy*, *precision*, dan *recall*.

1.2 Perumusan Masalah

Rumusan masalah dari penelitian ini adalah bagaimana menentukan fitur yang paling berpengaruh pada *dataset* penyakit jantung yang berasal dari UCI, serta penanganan data hilang menggunakan algoritma ReliefF dan menguji kinerja algoritma ReliefF pada penerapan *dataset* penyakit jantung dengan menggunakan *Backpropagation Neural Network* (BNN).

1.3 Pembatasan Masalah

Beberapa pembatasan masalah dalam penelitian ini adalah:

1. Penelitian ini hanya membahas pemilihan fitur terpenting dan penanganan data hilang pada *dataset* UCI yang memiliki dua label kelas yaitu sehat (*healthy*) dan sakit (*sick*). Selain itu, membahas bagaimana pengaruh masing-masing fitur pada kinerja klasifikasi, tapi tidak membahas *preprocessing* dan bagaimana data didapatkan.
2. Ukuran evaluasi kinerja pada penelitian ini menggunakan *accuracy*, *precision*, dan *recall*.

1.4 Tujuan

Tujuan dari penelitian ini adalah untuk menentukan fitur yang paling berpengaruh pada *dataset* penyakit jantung yang berasal dari UCI, serta

penanganan data hilang menggunakan algoritma ReliefF dan menguji kinerja algoritma ReliefF pada penerapan *dataset* penyakit jantung dengan menggunakan *Backpropagation Neural Network* (BNN).

1.5 Manfaat

Manfaat dari hasil penelitian ini adalah:

1. Mengetahui faktor yang paling berpengaruh dalam memprediksi penyakit jantung.
2. Sebagai masukan di bidang kesehatan untuk memprediksi penyakit jantung secara dini sehingga pihak medis dapat melakukan tindakan yang sesuai dalam penanganan pasien penyakit jantung.

DAFTAR PUSTAKA

- Al-barzinji, Y. M. S., Ahmad, M. A., & Saeed, B. K. (2020). *International Transaction Journal of Engineering , Management , & Applied Sciences & Technologies GENETIC DISTANCES AMONG EIGHT ORNAMENTAL*. 11(2), 1–10. <https://doi.org/10.14456/ITJEMAST.2020.294>
- Arora, N., & Kaur, P. D. (2020). A Bolasso based consistent feature selection enabled random forest classification algorithm: An application to credit risk assessment. *Applied Soft Computing Journal*, 86, 105936. <https://doi.org/10.1016/j.asoc.2019.105936>
- Baliarsingh, S. K., Ding, W., Vipsita, S., & Bakshi, S. (2019). A memetic algorithm using emperor penguin and social engineering optimization for medical data classification. *Applied Soft Computing Journal*, 85, 105773. <https://doi.org/10.1016/j.asoc.2019.105773>
- Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S., & Singh, P. (2021). Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning. *Computational Intelligence and Neuroscience*, 2021. <https://doi.org/10.1155/2021/8387680>
- Celikmih, K., Inan, O., & Uguz, H. (2020). Failure Prediction of Aircraft Equipment Using Machine Learning with a Hybrid Data Preparation Method. *Scientific Programming*, 2020. <https://doi.org/10.1155/2020/8616039>
- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers and Electrical Engineering*, 40(1), 16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- Cheema, J. R. (2014). A Review of Missing Data Handling Methods in Education Research. *Review of Educational Research*, 84(4), 487–508. <https://doi.org/10.3102/0034654314532697>
- Dai, B., Gu, H., Zhu, Y., Chen, S., & Rodriguez, E. F. (2020). On the Use of an Improved Artificial Fish Swarm Algorithm-Backpropagation Neural Network for Predicting Dam Deformation Behavior. *Complexity*, 2020. <https://doi.org/10.1155/2020/5463893>
- Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, 340–341, 250–261. <https://doi.org/10.1016/j.ins.2016.01.033>
- Desiani, A., Dewi, N. R., Fauza, A. N., Rachmatullah, N., Arhami, M., & Nawawi, M. (2021). Handling Missing Data Using Combination of Deletion Technique, Mean, Mode and Artificial Neural Network Imputation for Heart Disease Dataset. *Science and Technology Indonesia*, 6(4). <https://doi.org/10.26554/sti.2021.6.4.303-312>
- Drotár, P., Gazda, J., & Smékal, Z. (2015). An experimental comparison of

- feature selection methods on two-class biomedical datasets. *Computers in Biology and Medicine*, 66, 1–10. <https://doi.org/10.1016/j.compbiomed.2015.08.010>
- Furmanek, S., Mph, M. S., Glick, C., Chandler, T., Tella, M., Mattingly, W., & Ramirez, J. A. (2020). *University of Louisville Journal of Respiratory Infections The City of Louisville Encapsulates the United States Demographics*. 1–6. <https://doi.org/10.18297/jri/vol4/iss2/4.Abstract>
- Gore, S., & Govindaraju, V. (2016). *Theory and Relief Algorithm*. <https://doi.org/10.1007/978-3-319-19090-7>
- Hasnain, M., Pasha, M. F., Ghani, I., Imran, M., Alzahrani, M. Y., & Budiarto, R. (2020). Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking. *IEEE Access*, 8, 90847–90861. <https://doi.org/10.1109/ACCESS.2020.2994222>
- Jawad, J., Hawari, A. H., & Javaid Zaidi, S. (2021). Artificial neural network modeling of wastewater treatment and desalination using membrane processes: A review. *Chemical Engineering Journal*, 419(March), 129540. <https://doi.org/10.1016/j.cej.2021.129540>
- Johnson, T., Zhao, L., Manuel, G., Taylor, H., & Liu, D. (2019). Approaches to therapeutic angiogenesis for ischemic heart disease. *Journal of Molecular Medicine*, 97(2), 141–151. <https://doi.org/10.1007/s00109-018-1729-3>
- Lang, K. M., & Little, T. D. (2018). Principled missing data treatments. *Prevention Science*, 19(3), 284–294. <https://doi.org/10.1007/s11121-016-0644-5>
- Liu, Q., Xu, X., Tao, Y., & Wang, X. (2017). An Improved Decision Tree Method Base on RELIEFF for Medical Diagnosis. *Proceedings - 2016 International Conference on Digital Home, ICDH 2016*, 133–138. <https://doi.org/10.1109/ICDH.2016.037>
- Mhatre, T., & Varma, S. (2019). IJERT-Heart Disease Prediction using Evolutionary based Artificial Neural Network Heart Disease Prediction using Evolutionary based Artificial Neural Network. *International Journal of Engineering Research & Technology*, 8(08).
- Mishra, S., Vanli, O. A., Huffer, F. W., & Jung, S. (2016). Regularized discriminant analysis for multi-sensor decision fusion and damage detection with Lamb waves. *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2016*, 9803(850), 98032H. <https://doi.org/10.1117/12.2217959>
- Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access*, 7, 81542–81554. <https://doi.org/10.1109/ACCESS.2019.2923707>
- Mohd Nawi, N., Tosida, E. T., Hasbi, H., & Abdul Hamid, N. (2021). An Implementation of First and Second Order Neural Network Classification on

- Potential Drug Addict Repetition. *Emerging Advances in Integrated Technology*, 02(01), 18–29. <https://doi.org/10.30880/emait.2021.02.01.003>
- Oktay, A. A., Akturk, H. K., Esenboğa, K., Javed, F., Polin, N. M., & Jahangir, E. (2018). Pathophysiology and Prevention of Heart Disease in Diabetes Mellitus. *Current Problems in Cardiology*, 43(3), 68–110. <https://doi.org/10.1016/j.cpcardiol.2017.05.001>
- Osanaiye, O., Ogundile, O., Aina, F., & Periola, A. (2019). Feature selection for intrusion detection system in a cluster-based heterogeneous wireless sensor network. *Facta Universitatis - Series: Electronics and Energetics*, 32(2), 315–330. <https://doi.org/10.2298/fuee1902315o>
- Otto, C. M., Nishimura, R. A., Bonow, R. O., Carabello, B. A., Erwin, J. P., Gentile, F., Jneid, H., Krieger, E. V., Mack, M., McLeod, C., O’Gara, P. T., Rigolin, V. H., Sundt, T. M., Thompson, A., & Toly, C. (2021). 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Journal of the American College of Cardiology*, 77(4), e25–e197. <https://doi.org/10.1016/j.jacc.2020.11.018>
- Pedersen, A. B., Mikkelsen, E. M., Cronin-Fenton, D., Kristensen, N. R., Pham, T. M., Pedersen, L., & Petersen, I. (2017). Missing data and multiple imputation in clinical epidemiological research. *Clinical Epidemiology*, 9, 157–166. <https://doi.org/10.2147/CLEP.S129785>
- Pereira, R. B., Plastino, A., Zadrozny, B., & Merschmann, L. H. C. (2018). Categorizing feature selection methods for multi-label classification. *Artificial Intelligence Review*, 49(1), 57–78. <https://doi.org/10.1007/s10462-016-9516-4>
- Quaglio, M., Roberts, L., Bin Jaapar, M. S., Fraga, E. S., Dua, V., & Galvanin, F. (2020). An artificial neural network approach to recognise kinetic models from experimental data. *Computers and Chemical Engineering*, 135, 106759. <https://doi.org/10.1016/j.compchemeng.2020.106759>
- Rohman, E. M. S., & Rachmad, A. (2016). Predicting Medicine-Stocks by Using Multilayer Perceptron Backpropagation. *Journal Article IJSEA*, 5(3), 188–191. <https://ijsea.com/archive/volume5/issue3/IJSEA05031010.pdf>
- Rohrer, B. (2015). Methods for handling missing values. *Cortana Intelligence Gallery*. <https://gallery.cortanaintelligence.com/Experiment/Methods-for-handling-missing-values-1>
- Setti, S., & Wanto, A. (2019). Analysis of Backpropagation Algorithm in Predicting the Most Number of Internet Users in the World. *Jurnal Online Informatika*, 3(2), 110. <https://doi.org/10.15575/join.v3i2.205>
- Shah, S. M. S., Batool, S., Khan, I., Ashraf, M. U., Abbas, S. H., & Hussain, S. A. (2017). Feature extraction through parallel Probabilistic Principal Component Analysis for heart disease diagnosis. *Physica A: Statistical*

- Mechanics and Its Applications*, 482, 796–807.
<https://doi.org/10.1016/j.physa.2017.04.113>
- Shan, H., Xu, H., Zhu, S., & He, B. (2015). A novel channel selection method for optimal classification in different motor imagery BCI paradigms. *BioMedical Engineering Online*, 14(1), 1. <https://doi.org/10.1186/s12938-015-0087-4>
- Sisodia, D., & Sisodia, D. S. (2018). Prediction of Diabetes using Classification Algorithms. *Procedia Computer Science*, 132(Iccids), 1578–1585. <https://doi.org/10.1016/j.procs.2018.05.122>
- Sun, L., Kong, X., Xu, J., Xue, Z., Zhai, R., & Zhang, S. (2019). A Hybrid Gene Selection Method Based on ReliefF and Ant Colony Optimization Algorithm for Tumor Classification. *Scientific Reports*, 9(1), 1–14. <https://doi.org/10.1038/s41598-019-45223-x>
- Venkatesh, B., & Anuradha, J. (2019). A review of Feature Selection and its methods. *Cybernetics and Information Technologies*, 19(1), 3–26. <https://doi.org/10.2478/CAIT-2019-0001>
- Wang, Z., Zhang, Y., Chen, Z., Yang, H., Sun, Y., Kang, J., Yang, Y., & Liang, X. (2016). Application of ReliefF algorithm to selecting feature sets for classification of high resolution remote sensing image. *International Geoscience and Remote Sensing Symposium (IGARSS), 2016-Novem(41271013)*, 755–758. <https://doi.org/10.1109/IGARSS.2016.7729190>
- Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Springer*, 195, 21–30.
- Yang, H., & Garibaldi, J. M. (2015). A hybrid model for automatic identification of risk factors for heart disease. *Journal of Biomedical Informatics*, 58, S171–S182. <https://doi.org/10.1016/j.jbi.2015.09.006>
- Zainudin, M. N. S., Sulaiman, M. N., Mustapha, N., Perumal, T., Nazri, A. S. A., Mohamed, R., & Manaf, S. A. (2017). Feature selection optimization using hybrid relief-f with self-adaptive differential evolution. *International Journal of Intelligent Engineering and Systems*, 10(2), 21–29. <https://doi.org/10.22266/ijies2017.0430.03>
- Zajmi, L., Ahmed, F. Y. H., & Jaharadak, A. A. (2018). Concepts, Methods, and Performances of Particle Swarm Optimization, Backpropagation, and Neural Networks. *Applied Computational Intelligence and Soft Computing*, 2018. <https://doi.org/10.1155/2018/9547212>
- Zhang, B., Liang, X. L., Gao, H. Y., Ye, L. S., & Wang, Y. G. (2016). Models of logistic regression analysis, support vector machine, and back-propagation neural network based on serum tumor markers in colorectal cancer diagnosis. *Genetics and Molecular Research*, 15(2). <https://doi.org/10.4238/gmr.15028643>