

**PENANGANAN DATA HILANG
MENGUNAKAN SELEKSI FITUR C4.5 PADA KLASIFIKASI
PENYAKIT JANTUNG MENGGUNAKAN
*BACKPROPAGATION***

SKRIPSI

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

Oleh :

Dwi Fitrianti

NIM 08011281823023



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

2022

LEMBAR PENGESAHAN

**PENANGANAN DATA HILANG
MENGUNAKAN SELEKSI FITUR C4.5 PADA KLASIFIKASI
PENYAKIT JANTUNG MENGGUNAKAN
BACKPROPAGATION**

SKRIPSI

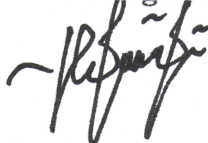
**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Sains Bidang Studi Matematika**

Oleh

DWI FITRIANTI

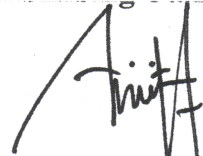
NIM. 08011281823023

Pembimbing Kedua



Dr. Yuli Andriani, M.Si
NIP. 197207021999032001

**Indralaya, April 2022
Pembimbing Utama**



Anita Desiani, S.Si., M.Kom
NIP. 197712112003122002

**Mengetahui,
Ketua Jurusan Matematika**



Drs. Sugandi Yahdin, M.M
NIP. 195807271986031003

PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Dwi Fitrianti
NIM : 08011281823023
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 penulis 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, 03 April 2022

Penulis



Dwi Fitrianti
NIM.08011281823023

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

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

Nama Mahasiswa : Dwi Fitrianti
NIM : 08011281823023
Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan Alam/Matematika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Sriwijaya “Penanganan Data Hilang Menggunakan Seleksi Fitur C4.5 Pada Klasifikasi Penyakit Jantung Menggunakan *Backpropagation*”. Dengan hak bebas royalti 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, 03 April 2022

Penulis



Dwi Fitrianti
NIM.08011281823023

LEMBAR PERSEMBAHAN

Kupersembahkan skripsi ini untuk:

Allah Subhanahu Wa Ta'ala,

Kedua orang tuaku tersayang,

Nenek, Kakak, Ayuk, Adik, dan Keponakanku,

Keluarga besarku,

Semua guru dan dosenku,

Sahabat-sahabatku,

Almamaterku

Moto

“Everything’s gonna be okay.”

KATA PENGANTAR

Puji syukur kehadirat Allah Subhanahu wa Ta'ala yang telah memberikan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penanganan data hilang menggunakan seleksi fitur c4.5 pada klasifikasi penyakit jantung menggunakan *Backpropagation*” 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 tersayang, **Marsilis dan Aidid**, yang selalu menyayangi, mendidik, menasehati, membimbing, mendukung dan terus mendoakan anaknya. Terima kasih atas segala pengorbanan dan hal luar biasa yang telah diberikan hingga detik ini dan sampai kapanpun.
2. Bapak **Drs. Sugandi Yahdin, M.M** selaku Ketua Jurusan Matematika FMIPA Universitas Sriwijaya yang telah memberikan motivasi, semangat, serta pengarahan kepada penulis selama proses perkuliahan.
3. Ibu **Dr. Dian Cahyawati Sukanda, M.Si** selaku Sekretaris Jurusan Matematika FMIPA Universitas Sriwijaya yang telah memberi arahan mengenai urusan akademik kepada penulis.
4. Ibu **Anita Desiani, S.Si., M.Kom** selaku dosen pembimbing utama yang telah bersedia mencurahkan waktu, pikiran, dan juga tenaga dalam memberikan bimbingan serta pengarahan baik selama masa kuliah hingga

masa skripsi, dan selalu memberikan motivasi dan didikan dalam hal kompetisi dan semangat untuk berprestasi.

5. Ibu **Dr. Yuli Andriani, M.Si** selaku dosen pembimbing pendamping yang telah bersedia mencurahkan waktu, pikiran, dan tenaga dalam memberikan pengarahan serta bimbingan selama proses pembuatan skripsi hingga sidang sarjana dengan pengertian.
6. Bapak **Drs. Endro Setyo C, M.Si** dan Ibu **Sisca Octarina, M.Sc** selaku dosen pembahas dan penguji yang telah memberikan pembelajaran dalam pembuatan skripsi melalui kritik, tanggapan, dan saran yang sangatlah bermanfaat dalam penulisan skripsi.
7. Ibu **Dr. Yulia Resti, M.Si**, selaku dosen pembimbing akademik yang telah memberikan bimbingan dan arahan dalam urusan akademik penulis.
8. **Seluruh Dosen di Jurusan Matematika FMIPA** yang telah memberikan banyak sekali pembelajaran dalam bidang akademik maupun non akademik, nasihat, dan bimbingan selama masa perkuliahan.
9. Pak **Irwansyah** selaku admin jurusan dan Ibu **Hamidah** selaku pegawai tata usaha Jurusan Matematika FMIPA yang telah banyak sekali membantu penulis dalam urusan administrasi selama masa perkuliahan.
10. **Seluruh guru** yang telah memberikan ilmu, nasihat, serta motivasi yang sangatlah bermanfaat hingga dapat mengantarkan penulis pada pendidikan ini.
11. Mamak, Kakak, Adik, Ayuk, dan Keponakanku, **Sukarti, Ady Saputera, Gilang Pandu Satria, Dewi, dan Galuh**, yang senantiasa selalu menyayangi,

mendoakan, memberi nasihat dan motivasi yang sangatlah berharga, beserta keluarga besar yang selalu mendukung penulis.

12. Sahabat seperjuangan dibangku perkuliahan, **Sekar, Henny, Nisa, Irma, Muti, Dita, Putri, Mar, Azwar, Wahyu, Tim Fighting, Tim BrotherSquad, Tim Pengen Anak Ucul, Tim Bimbingan Bu Anita dan Tim PP Club**. Terima kasih sudah menjadi orang-orang hebat yang selalu membantu dengan tulus, dan memberi dukungan lewat energi positif.
13. **Tim PHP2D (Nisa, Henny, Muti, Rifa, Itak, Aga, Miwa, Tasya, Azwar, Wahyu, Kak Indri, Kak Abu, Kak yogi, Kak Mega, dan Kak Olin)** yang telah menjadi keluarga dari Beti hingga sampai sekarang ini.
14. **Keluarga Matematika 2018, BPH Himastik Beraksi dan Gelora Karya, Tim PHP2D, dan rekan-rekan** selama perkuliahan.
15. Sahabat-sahabatku sewaktu TK, SD, SMP, SMA, serta Kakak-kakak tingkat angkatan 2016 dan 2017 serta adik-adik tingkat angkatan 2019 dan 2020, terima kasih atas segala kebaikan dan bantuan.
16. Seluruh pihak yang telah membantu penulis dan tidak dapat dituliskan satu per satu. Semoga semua bantuan dan kebaikan yang telah diberikan, mendapatkan pahala dari Allah.

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, April 2022

Penulis

**HANDLING OF MISSING VALUE
USING THE FEATURE SELECTION C4.5 IN CLASSIFICATION
OF HEART DISEASE USING BACKPROPAGATION**

By:

**Dwi Fitrianti
08011281823023**

ABSTRACT

Missing Value can result in weak prediction accuracy, so it needs to be addressed. The way to deal with lost data is to use the deletion method. This method requires a classification algorithm to select the most influential features. In this study, the University of California Irvine (UCI) heart attack dataset was used which has a weakness, namely it has missing data on several features. This study uses the C4.5 Algorithm to overcome missing data by selecting important features and using Backpropagation for testing. The test results obtained an accuracy value of 78.9116 % for testing without the application of the C4.5 algorithm on missing data, while the accuracy value of 80.6112% for testing with the application of the C4.5 algorithm on missing data using 10 features in the test. It can be concluded that the application of the C4.5 algorithm to select influential features and the use of Backpropagation for testing can improve accuracy.

Keywords : *C4.5 Algorithm, Missing Value, Backproppagation, and Accuracy.*

**PENANGANAN DATA HILANG
MENGUNAKAN SELEKSI FITUR C4.5 PADA KLASIFIKASI
PENYAKIT JANTUNG MENGGUNAKAN
*BACKPROPAGATION***

Oleh:

**Dwi Fitrianti
08011281823023**

ABSTRAK

Kehilangan data dapat mengakibatkan lemahnya keakuratan suatu prediksi, sehingga perlu diatasi. Cara untuk mengatasi data hilang digunakan metode *deletion*. Metode ini memerlukan algoritma klasifikasi untuk menyeleksi fitur yang paling berpengaruh. Pada penelitian ini digunakan *dataset* serangan jantung *Univercity of California Irvine* (UCI) yang memiliki kelemahan yaitu memiliki data hilang pada beberapa fitur. Penelitian ini menggunakan Algoritma C4.5 untuk mengatasi data hilang dengan menyeleksi fitur-fitur yang penting, dan pada pengujian digunakan *Backpropagation*. Hasil pengujian diperoleh nilai akurasi sebesar 78,9116 % untuk pengujian tanpa penerapan algoritma C4.5 pada data hilang, sedangkan nilai akurasi sebesar 80,6112% untuk pengujian dengan penerapan algoritma C4.5 pada data hilang dengan digunakan 10 fitur dalam pengujian. Dapat disimpulkan bahwa penerapan algoritma C4.5 untuk menyeleksi fitur yang berpengaruh dan penggunaan *Backpropagation* untuk pengujian dapat meningkatkan akurasi.

Kata Kunci : *Algoritma C4.5, Data Hilang, Backpropagation, dan Akurasi*

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSEMBAHAN	iii
KATA PENGANTAR.....	iv
ABSTRACT	vii
ABSTRAK	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Pembatasan Masalah	4
1.4 Tujuan	5
1.5 Manfaat	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penyakit Jantung	6
2.2 Data Hilang	6
2.3 Klasifikasi.....	7
2.4 <i>Preprocessing</i>	7
2.5 <i>Decision Tree</i>	8
2.7 Evaluasi Metode.....	10
2.7.1. <i>Backpropagation</i>	10
2.7.2. <i>Confusion Matrix</i>	13
BAB III METODOLOGI PENELITIAN	17
3.1 Tempat.....	17
3.2 Waktu	17
3.3 Alat	17
3.4 Tahap Penelitian.....	17
BAB IV HASIL DAN PEMBAHASAN.....	20
4.1 Deskripsi Data	20
4.2 Perhitungan Manual Penerapan Algoritma untuk Seleksi Fitur dan Klasifikasi Menggunakan <i>Backpropagation</i>	23
4.2.1. Penerapan Algoritma C4.5 dengan Perhitungan Manual.....	23
4.2.2 Penerapan Klasifikasi Menggunakan <i>Backpropagation</i> dengan Perhitungan Manual.....	30
4.3 Penerapan Algoritma C4.5 dan <i>Backpropagation</i> pada Data Penelitian	36
4.3.1. Algoritma C4.5 untuk data penelitian.....	36
4.3.2. Penerapan <i>Backpropagation</i> pada Data Penelitian	37
4.4 Analisis Hasil	47

BAB V KESIMPULAN DAN SARAN	48
5.1 Kesimpulan.....	48
5.2 Saran.....	48
DAFTAR PUSTAKA	49
LAMPIRAN.....	53

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	14
Tabel 2.2 Kategori Nilai Akurasi	15
Tabel 4.1 <i>Dataset</i> Serangan Jantung.....	20
Tabel 4.2 Keterangan Fitur	21
Tabel 4.3 Data Contoh untuk Perhitungan Manual.....	23
Tabel 4.4 Nilai <i>Entropy</i> dan <i>Gain</i> dari 14 fitur.....	26
Tabel 4.5 Nilai <i>Entropy</i> dan <i>Gain</i> dari <i>Oldpeak</i> - <i>Ya</i>	28
Tabel 4.6 Bobot Input Pelatihan Contoh Perhitungan Manual	31
Tabel 4.7 Bobot Masukan Pelatihan Contoh Perhitungan Manual dari <i>Hidden Layer</i> Menuju <i>Output Layer</i>	32
Tabel 4.8 Bobot Terbaru pada <i>Output Layer</i> dan <i>Hidden Layer</i>	36
Tabel 4.9 Fitur yang Terdapat pada Pohon Keputusan	37
Tabel 4.10 Hasil Perbandingan Nilai Akurasi pada Setiap Seleksi Fitur.....	38
Tabel 4.11 Fitur-Fitur pada Diagnosa Penyakit Jantung.....	39
Tabel 4.12 <i>Confusion Matrix</i> Metode <i>Backpropagation</i> pada Data Sebelum Penerapan Algoritma C4.5	40
Tabel 4.13 Hasil Evaluasi Data Sebelum Penerapan C4.5.....	42
Tabel 4.14 Fitur-Fitur pada Diagnosa Penyakit Jantung.....	43
Tabel 4.15 <i>Confusion Matrix</i> Metode <i>Backpropagation</i> pada Data Setelah Penerapan Algoritma C4.5	43
Tabel 4.16 Hasil Evaluasi Data dengan Metode <i>Backpropagation</i> Setelah Penerapan C4.5.....	46
Tabel 4.17 Perbandingan Hasil Akurasi Prediksi Penyakit Jantung	47
Tabel 4.18 Perbandingan Hasil Penelitian dengan Penelitian Lain	48

DAFTAR GAMBAR

Gambar 2.1 Struktur Pohon Keputusan	8
Gambar 2.2 Arsitektur <i>Backpropagation</i>	11
Gambar 4.1 Akar Pohon Keputusan Perhitungan Manual	28
Gambar 4.2 Pohon Keputusan C4.5 Contoh Perhitungan Manual.....	30
Gambar 4.3 Pohon Keputusan Algoritma C4.5	37

BAB I

PENDAHULUAN

1.1 Latar Belakang

Klasifikasi telah banyak digunakan dalam berbagai bidang ilmu pengetahuan. Klasifikasi merupakan proses dikelompokkannya objek ke dalam sebuah kelas yang sesuai dengan objek tersebut (Tharwat, 2018). Klasifikasi dapat menggunakan beberapa metode antara lain algoritma C4.5 dan *Backpropagation*. Algoritma C4.5 adalah salah satu algoritma klasifikasi yang paling sering digunakan dalam melakukan pembuatan keputusan, kinerjanya dalam membuat pohon keputusan dan memprediksi hasil terbaik memakan waktu yang minimum, serta tingkat akurasinya juga tinggi (Muslim *et al.*, 2018). Algoritma C4.5 memiliki beberapa kelebihan yaitu mengatasi *missing value*, mampu menangani atribut dengan kontinu, memangkas pohon keputusan untuk mengatasi *overfitting*, dan algoritma C4.5 bisa mempersempit keputusan yang dihasilkan pohon untuk mengurangi ukurannya dan mengoptimalkan jalur keputusan (Cherfi *et al.*, 2018).

Beberapa penelitian menggunakan algoritma C4.5 antaranya dilakukan oleh Suyatno, Nhita & Rohmawati, (2018) melakukan penelitian dengan menerapkan algoritma C4.5 terhadap perkiraan curah hujan di kabupaten Bandung, dengan mendapatkan hasil akurasi sebelum pemangkasan fitur menggunakan algoritma C4.5 sebesar 60% dan setelah pemangkasan fitur akurasi menjadi sebesar 93,33%. Penelitian prediksi pasien diabetes rawat inap di rumah sakit menggunakan algoritma C4.5 dengan mendapatkan hasil akurasi sebesar 82,74%, presisi 87,1 %

dan sensitivitas 82,7% (Pujianto *et al.*, 2019).

Prasetyo & Prasetyo, (2020) melakukan penelitian terhadap diagnosa penyakit jantung menggunakan algoritma C4.5 dan didapatkan akurasi sebesar 72,98% dan kemudian dikombinasikan dengan *Bagging* dan mendapat akurasi sebesar 81,84%. Algoritma C4.5 bisa mengelola semua tipe data, namun lebih diutamakan data katagorik (Sulistiani & Aldino, 2020), sedangkan salah satu metode klasifikasi yang bisa mengelola semua jenis data adalah *Backpropagation* (Lillicrap *et al.*, 2020).

Backpropagation merupakan metode yang dapat memberikan hasil akurat dalam melakukan suatu prediksi, karena metode ini melakukan pelatihan yang berulang-ulang untuk mendapatkan model yang terbaik dan juga dapat dianalisa secara matematika (Siregar & Wanto, 2017). *Backpropagation* memiliki tujuan awal yaitu meminimalkan jumlah *error* pada *output*, secara bertahap dengan menghasilkan bobot baru (Hameed *et al.*, 2016). Beberapa penelitian yang menggunakan *Backpropagation* diantaranya adalah penelitian yang dilakukan oleh Ismael & Abdel-Qader, (2018), diperoleh nilai akurasi sebesar 91,9% dalam klasifikasi penyakit tumor otak. Nagra *et al.*, (2020) melakukan penelitian dengan metode yang sama menggunakan data diagnosa penyakit ginjal kronis didapatkan akurasi sebesar 94,8%, sehingga *Backpropagation* dapat digunakan untuk mengklasifikasi penyakit jantung (Ramakrishnan & Sahaya, 2015).

Penyakit jantung adalah penyakit yang disebabkan oleh penyempitan pada pembuluh darah yang terjadi karena penumpukan plak, dimana terdapat gangguan pada sistem pembuluh darah besar sehingga menyebabkan jantung dan peredaran

darah tidak berkerja dengan baik (Elfi *et al.*, 2019). Di dunia Penyakit jantung masih menjadi penyakit penyebab kematian nomor satu (Mohan *et al.*, 2019). Serangan jantung dapat terrjadi secara tiba-tiba dan dapat berakibat pada kondisi gagal jantung, sehingga sangat diperlukan untuk mendiagnosa secara dini dan akurat sebagai pencegahan penyakit jantung (Haq *et al.*, 2018).

Pada riset penyakit jantung, data yang sering digunakan adalah data dari *Univercity of California Irvine (UCI) Machine Learning Repository*. Pada *dataset* ini terdapat 76 fitur untuk mendiagnosa penyakit jantung, dari 76 fitur tersebut yang hanya dipublikasi oleh UCI adalah 14 fitur yaitu *Age*, *Sex*, *Cp*, kolesterol, kadar gula, angina induksi, elektrokardiografi, *flaurosopy*, tekanan darah, detak jantung maksimal, *slope*, *oldpeak*, atau kemiringan ST, denyut jantung dan fitur *healthy* (sehat) dan *sick* (sakit) yang terdapat pada fitur label. Berdasarkan fitur-fitur tersebut terdapat data hilang yang menyebabkan kurang akuratnya dalam mendiagnosa penyakit jantung.

Penanganan data hilang dapat dilakukan dengan beberapa metode, yaitu membuang fitur (*deletion*), pendugaan nilai baru, dan prediksi. Pada pendugaan nilai baru, dapat dilakukan pengisian nilai baru dengan menggunakan *mean*, modus, median. Pada metode *deletion* terdapat dua cara, yaitu dengan menghapus fitur secara langsung atau dengan menyeleksi fitur-fitur yang mempengaruhi menggunakan metode algoritma klasifikasi (Desiani *et al.*, 2021). Penseleksian fitur-fitur yang berpengaruh pada sebuah data dapat digunakan algoritma C4.5. Penelitian yang dilakukan oleh Mutrofin *et al.* (2019) menggunakan algoritma C4.5 untuk menyeleksi fitur-fitur yang paling berpengaruh pada data prediksi

siswa terkena *dropout* untuk mengatasi data hilang. Nagra *et al.*, (2020) menggunakan algoritma C4.5 untuk menyeleksi fitur-fitur yang paling berpengaruh pada data diagnosa kesalahan bilah turbin angin.

Penelitian ini mengkombinasikan algoritma C4.5 dan *Backpropagation*, dimana algoritma C4.5 digunakan untuk menyeleksi fitur-fitur yang tidak terlalu berpengaruh dalam diagnosa penyakit jantung dalam mengatasi data hilang. Kemudian, hasil dari seleksi fitur-fitur dengan C4.5 akan diukur kinerjanya menggunakan *Backpropagation* dengan mengukur akurasi, presisi, dan sensitivitas untuk mengetahui sejauh mana C4.5 dan *Backpropagation* dapat menghasilkan performa yang baik pada *dataset* diagnosa penyakit jantung.

1.2 Perumusan Masalah

Rumusan masalah dari penelitian ini adalah bagaimana mencari fitur paling berpengaruh pada data diagnosa penyakit jantung menggunakan algoritma C4.5 dalam mengatasi data hilang, dan mengukur hasil kinerja klasifikasi menggunakan *Backpropagation* dengan melihat nilai akurasi, presisi, dan sensitivitas.

1.3 Pembatasan Masalah

Beberapa pembatasan masalah dalam penelitian ini adalah:

1. Penelitian ini hanya membahas 14 fitur yang digunakan untuk diagnosa penyakit jantung, diantaranya adalah *Age*, *Sex*, *Cp*, *Chol*, *fbs*, angina induksi, elektrokardiografi, *flaurosopy*, *tresrbps*, *thalac*, *oldpeak*, *slope* atau kemiringan ST, denyut jantung, sehat dan sakit yang digunakan dalam data tersebut.

2. Ukuran evaluasi kinerja pada penelitian ini menggunakan akurasi, presisi, dan sensitivitas.

1.4 Tujuan

Tujuan dari penelitian ini adalah untuk mencari fitur paling berpengaruh pada data diagnosa penyakit jantung menggunakan algoritma C4.5 dalam mengatasi data hilang, dan mengukur hasil kinerja klasifikasi menggunakan *Backpropagation* dengan melihat nilai akurasi, presisi, dan sensitivitas.

1.5 Manfaat

1. Penerapan algoritma C4.5 dapat mengetahui fitur yang paling berpengaruh dalam diagnosa penyakit jantung.
2. Dapat mengetahui alternatif metode mengatasi data hilang.

DAFTAR PUSTAKA

- Cherfi, A., Nouira, K., & Ferchichi, A. (2018). Very fast C4.5 decision tree algorithm. *Applied Artificial Intelligence*, 32(2), 119–137. <https://doi.org/10.1080/08839514.2018.1447479>
- Dalen, J. E., Alpert, J. S., Goldberg, R. J., & Weinstein, R. S. (2014). The epidemic of the 20th century: coronary heart disease. *American Journal of Medicine*, 127(9), 807–812. <https://doi.org/10.1016/j.amjmed.2014.04.015>
- 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), 303–312. <https://doi.org/10.26554/sti.2021.6.4.303-312>
- Elfi, E. F., Ilhami, Y. R., & Darwin, E. (2019). The role of endothelial microparticle in coronary heart disease as the complications of diabetes mellitus. *Jurnal Biodjati*, 4(1), 31–39. <https://doi.org/10.15575/biodjati.v4i1.4164>
- Hameed, A. A., Karlik, B., & Salman, M. S. (2016). Back-propagation algorithm with variable adaptive momentum. *Knowledge-Based Systems*, 114, 79–87. <https://doi.org/10.1016/j.knosys.2016.10.001>
- Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., Sun, R., & Garcíá-Magarinõ, I. (2018). A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. *Mobile Information Systems*, 2018. <https://doi.org/10.1155/2018/3860146>
- 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>
- Ismael, M. R., & Abdel-Qader, I. (2018). Brain tumor classification via statistical features and back-propagation neural network. *IEEE International Conference on Electro Information Technology*, 2018-May, 252–257. <https://doi.org/10.1109/EIT.2018.8500308>
- Izhari, F., Dhany, H. W., Zarlis, M., & Sutarman. (2018). Analysis backpropagation methods with neural network for prediction of children's ability in psychomotoric. *Journal of Physics: Conference Series*, 978(1). <https://doi.org/10.1088/1742-6596/978/1/012085>

- Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, 21(6), 335–346. <https://doi.org/10.1038/s41583-020-0277-3>
- 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. *Proc. of SPIE*, 9803. <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>
- Muslim, M. A., Rukmana, S. H., Sugiharti, E., Prasetyo, B., & Alimah, S. (2018). Optimization of C4.5 algorithm-based particle swarm optimization for breast cancer diagnosis. *Journal of Physics: Conference Series*, 983(1). <https://doi.org/10.1088/1742-6596/983/1/012063>
- Mutrofin, S., Khalimi, A. M., Kurniawan, E., Ginardi, R. V. H., Fatichah, C., & Sari, Y. A. (2019). Detection of potentially students drop out of college in case of missing value using C4.5. *ICSECC 2019 - International Conference on Sustainable Engineering and Creative Computing: New Idea, New Innovation, Proceedings*, 349–354. <https://doi.org/10.1109/ICSECC.2019.8907014>
- Nagra, A. A., Han, F., Ling, Q. H., Abubaker, M., Ahmad, F., Mehta, S., & Apasiba, A. T. (2020). Hybrid self-inertia weight adaptive particle swarm optimisation with local search using C4.5 decision tree classifier for feature selection problems. *Connection Science*, 32(1), 16–36. <https://doi.org/10.1080/09540091.2019.1609419>
- Nurmasani, A., & Pristyanto, Y. (2021). Algoritme stacking untuk klasifikasi penyakit jantung pada dataset imbalanced class. *Jurnal Pseudocode*, 8(1), 21–26. <https://doi.org/10.33369/pseudocode.8.1.21-26>
- Permana, B. A. C., Ahmad, R., Bahtiar, H., Sudianto, A., & Gunawan, I. (2021). Classification of diabetes disease using decision tree algorithm (C4.5). *Journal of Physics: Conference Series*, 1869(1). <https://doi.org/10.1088/1742-6596/1869/1/012082>
- Prasetyo, E., & Prasetyo, B. (2020). Increased classification accuracy C4 . 5 algorithm using bagging techniques in diagnosing heart disease. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 7(5), 1035–1040. <https://doi.org/10.25126/jtiik.202072379>
- Priyam, A., Gupta, R., Rathee, A., & Srivastava, S. (2013). Comparative analysis of decision tree classification algorithms. *International Journal of Current Engineering and Technology*, 3(2), 334–337.
- Pujianto, U., Setiawan, A. L., Rosyid, H. A., & Salah, A. M. M. (2019).

- Comparison of naïve bayes algorithm and decision tree C4.5 for hospital readmission diabetes patients using HbA1c Measurement. *Knowledge Engineering and Data Science*, 2(2), 58. <https://doi.org/10.17977/um018v2i22019p58-71>
- Putra, P. D., & Rini, D. P. (2019). Prediksi penyakit jantung dengan algoritma klasifikasi. *Prosiding Annual Research Seminar 2019*, 5(1), 978–979.
- Raja, P. S., & Thangavel, K. (2020). Missing value imputation using unsupervised machine learning techniques. In *Soft Computing* (Vol. 24, Issue 6). Springer Berlin Heidelberg. <https://doi.org/10.1007/s00500-019-04199-6>
- Ramakrishnan, M., & Sahaya, A. N. A. (2015). Groundnut leaf disease detection and classification by using back propagation algorithm. *2015 International Conference on Communication and Signal Processing, ICCSP 2015, 7092512506*, 964–968. <https://doi.org/10.1109/ICCSP.2015.7322641>
- Ryu, S., Kim, M., & Kim, H. (2020). Denoising autoencoder-based missing value imputation for smart meters. *IEEE Access*, 8, 40656–40666. <https://doi.org/10.1109/ACCESS.2020.2976500>
- Santra, A. K., & Christy, C. J. (2012). Genetic algorithm and confusion matrix for document clustering. *International Journal of Computer Science Issues*, 9(1), 322–328.
- Siregar, S. P., & Wanto, A. (2017). Analysis of artificial neural network accuracy using backpropagation algorithm in predicting process (forecasting). *International Journal Of Information System & Technology*, 1(1), 34. <https://doi.org/10.30645/ijistech.v1i1.4>
- Somvanshi, M., Chavan, P., Tambade, S., & Shinde, S. V. (2017). A review of machine learning techniques using decision tree and support vector machine. *Proceedings - 2nd International Conference on Computing, Communication, Control and Automation, ICCUBEA 2016*. <https://doi.org/10.1109/ICCUBEA.2016.7860040>
- Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, 27(2), 130–135. <https://doi.org/10.11919/j.issn.1002-0829.215044>
- Sugianto, C. A. (2017). Penerapan teknik data mining untuk menentukan ahsil seleksi asuk Sman 1 Gibeber untuk siswa baru menggunakan decision treem. 39–43. <https://doi.org/10.31227/osf.io/vedu7>
- Sulistiani, H., & Aldino, A. A. (2020). Decision tree C4.5 algorithm for tuition aid grant program classification (Case study: department of information system, Universitas Teknokrat Indonesia). *Jurnal Ilmiah Edutic*, 7(1), 40–50. <https://doi.org/10.21107/edutic.v7i1.8849>

- Suyatno, J. A., Nhita, F., & Rohmawati, A. A. (2018). Rainfall forecasting in Bandung regency using C4.5 algorithm. *2018 6th International Conference on Information and Communication Technology, ICoICT 2018*, 0(c), 324–328. <https://doi.org/10.1109/ICoICT.2018.8528725>
- Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*, 17(1), 168–192. <https://doi.org/10.1016/j.aci.2018.08.003>
- Yahdin, S., Desiani, A., Amran, A., Rodiah, D., & Solehan. (2019). Pattern recognition for study period of student in Mathematics Department with C4.5 algorithm data mining technique at the faculty of mathematics and natural science Universitas Sriwijaya. *Journal of Physics: Conference Series*, 1282(1). <https://doi.org/10.1088/1742-6596/1282/1/012014>