



Classification of Heart Failure Disease Using Supervised Learning

Klasifikasi Penyakit Gagal Jantung Menggunakan *Supervised Learning*

Ena Tasia^{1*}, Raja Zaid Ibnu Zarier Ismail², Septi Kenia Pita Loka³,
Yulia Ikhsani⁴, Rahweni Ocviani⁵

^{1,2,3,4,5}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Sultan Syarif Kasim Riau

E-Mail: ¹12050321658@students.uin-suska.ac.id,
²12050310503@students.uin-suska.ac.id, ³12050323230@students.uin-suska.ac.id,
⁴12050320529@students.uin-suska.ac.id, ⁵11950320049@students.uin-suska.ac.id

Corresponding Author: Ena Tasia

Abstract

The heart is a vital organ in humans that has a very important role in the human body so that a person suffering from heart disease is likely to experience complaints during his lifetime. Heart failure is a major health problem worldwide. In 2013, the prevalence of heart failure shown by the Basic Health Research (Riskesdas) in Indonesia was estimated to be around 0.3% or around 530,068 people. With so many people affected by heart failure disease, researchers are interested in classifying to predict the possibility of heart failure disease being the cause of death. This time the researchers will use 3 algorithms, namely the c4.5 algorithm, Support-Vector Machine (SVM), and Naïve Bayes Classifier (NBC) with data sharing techniques using haldout. The dataset used has 13 attributes and 299 data, obtained as much as 239 training data and 60 testing data. The results of this study are in the form of classification results using Python tools with the acquisition of the C.45 algorithm accuracy value of 80%, the NBC algorithm of 68% while the SVM algorithm is 68%. It can be seen that the c4.5 algorithm is more optimal for the classification of patient data affected by heart failure with an accuracy value of 80%.

Keyword: C4.5, Classification, Heart Failure, Holdout, NBC, SVM

Abstrak

Jantung adalah organ vital pada manusia yang memiliki peran sangat penting pada tubuh manusia sehingga seseorang yang menderita penyakit jantung berkemungkinan mengalami keluhan semasa hidupnya. Di seluruh dunia gagal jantung (*Heart Failure*) merupakan masalah kesehatan utama. Di tahun 2013 prevalensi gagal jantung yang ditunjukkan oleh Riset Kesehatan Dasar (Riskesdas) Indonesia diperkirakan sekitar 0,3% atau sekitar 530.068 orang. Dengan banyaknya orang yang terkena penyakit gagal jantung maka peneliti tertarik untuk melakukan klasifikasi untuk memprediksi kemungkinan penyakit gagal jantung menjadi penyebab kematian. Kali ini peneliti akan menggunakan 3 algoritma yaitu algoritma c4.5, *Support-Vektor Machine* (SVM), dan *Naïve Bayes Classifier* (NBC) dengan teknik pembagian data menggunakan *haldout*. Dataset yang digunakan memiliki 13 atribut dan 299 data, diperoleh sebanyak 239 data *training* dan 60 data *testing*. Hasil dari penelitian ini berupa hasil klasifikasi menggunakan *tools Python* dengan perolehan nilai akurasi algoritma C.45 sebesar 80%, algoritma NBC sebesar 68% sedangkan algoritma SVM sebesar 68%. Dapat diketahui bahwa algoritma c4.5 lebih optimal untuk klasifikasi data pasien yang terkena gagal jantung dengan nilai akurasi 80%.

Kata Kunci: C4.5, Gagal Jantung, Holdout, Klasifikasi, NBC, SVM

1. PENDAHULUAN

Bagi setiap makhluk hidup kesehatan adalah hal yang paling penting. Kesehatan juga menjadi pendorong utama peningkatan kualitas sumber daya manusia, pengentasan kemiskinan dan pembangunan

ekonomi. Sebagai orang yang memiliki tujuan dan pencapaian dalam hidup, seorang individu harus selalu berusaha untuk menjadi sehat secara fisik dan produktif [1]. Karena ketika kondisi tubuh tidak dalam keadaan yang baik, maka akan timbul berbagai macam penyakit yang salah satunya penyakit *kardiovaskuler* yang menyerang organ jantung [2].

Salah satu organ vital pada manusia adalah jantung, yang mana bertugas untuk memompa darah ke seluruh tubuh. Hal inilah yang membuat jantung mengambil peran sangat penting dalam tubuh manusia, sehingga penderita penyakit jantung kemungkinan besar akan mengalami ketidaknyamanan sepanjang hidupnya. [3]. Penyakit jantung merupakan penghambat fungsi kerja jantung. Penyakit atau kondisi seperti *kardiovaskular*, penyakit jantung koroner, dan gagal jantung [4].

Gagal jantung (Heart Failure) adalah kasus kesehatan primer pada semua dunia, menggunakan prevalensi lebih menurut 5,8 juta pada Amerika Serikat & lebih menurut 23 juta pada semua dunia [5]. Di tahun 2013 prevalensi gagal jantung yang ditunjukkan oleh Riset Kesehatan Dasar (Riskesdas) Indonesia diperkirakan sekitar 0,3% atau sekitar 530.068 orang [6]. Diperkirakan 5% menurut pasien yang dirawat di rumah sakit. 4,7% wanita dan 5,1% pria. Insiden tahunan gagal jantung diperkirakan 2,3-3,7 per seribu pasien tiap tahun. Angka kejadian gagal jantung diperkirakan akan semakin tinggi di masa depan karena meningkatnya harapan hidup dan perkembangan terapi untuk infark miokard, yang mengarah pada pemulihan harapan hidup pasien dengan melemahnya fungsi jantung [7]. Dengan banyaknya orang yang terkena penyakit gagal jantung maka dibutuhkan klasifikasi untuk memprediksi kemungkinan penyakit gagal jantung menjadi penyebab kematian dengan suatu pengolahan data yaitu dengan metode data mining.

Data mining adalah proses pemilahan warta berdasarkan seperangkat data yang sangat besar melalui penerapan prosedur pemecahan dan teknik penarikan pada bidang statistik, pembelajaran mesin dan sistem manajemen basis data [8]. Klasifikasi adalah metode untuk menentukan anggota suatu kelas yang mana kelasnya telah ditentukan terlebih dahulu berdasarkan karakter data yang sama [9].

Sepharni, dkk pada tahun 2022 tentang klasifikasi penyakit jantung menggunakan algoritma C4.5 menghasilkan akurasi sebesar 79% [10]. Ramadhan & Khoirunnisa pada tahun 2021 menggunakan *Support Vector Machine* (SVM) untuk melakukan klasifikasi Malaria dengan akurasi sebesar 92,3% [11]. Irfan dkk. pada tahun 2019 yang membandingkan algoritma C4.5 dengan *Support Vector Machine* (SVM) untuk identifikasi hama dan penyakit pada tanaman cabai, algoritma C4.5 lebih akurat dibandingkan *Support Vector Machine* (SVM) dengan akurasi 89,29%, sedangkan algoritma SVM menghasilkan akurasi 82,33% [12]. A`yuniyah dkk pada tahun 2022 mengimplementasikan *Naïve Bayes Classifier* (NBC) pada penyakit ginjal kronik dengan nilai akurasi mencapai 96,43% [13]. Desiani dkk. pada tahun 2022 mengimplementasikan *Naïve Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM) pada penyakit kardiovaskular dengan nilai akurasi tertinggi dari NBC sebesar 71% menggunakan *K-Fold Cross Validation* [14]. Berdasarkan penelitian tersebut, maka dilakukan perbandingan algoritma C4.5, *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier* (NBC) untuk mengetahui algoritma mana yang terbaik dalam memprediksi penyakit jantung dengan menggunakan *Hold Out* sebagai metode pembagian data.

Berdasarkan permasalahan yang telah dijelaskan pada pendahuluan, maka penelitian ini akan melakukan perbandingan klasifikasi gagal jantung dengan menggunakan tiga algoritma untuk mengetahui manakah yang memiliki performa atau akurasi terbaik pada pengujian data gagal jantung. Oleh karena itu pada penelitian ini diangkatlah sebuah judul Metode Klasifikasi Supervised Learning pada Penyakit Gagal Jantung.

2. BAHAN DAN METODE

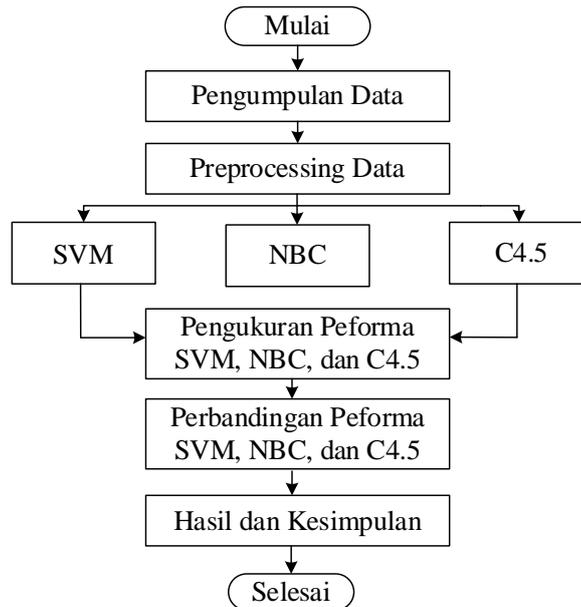
Tahapan yang digunakan untuk penelitian ini terbagi menjadi lima tahap yakni pengumpulan data, pre-processing data, proses klusterisasi dengan algoritma, pengukuran validitas cluster serta analisis dan hasil algoritma terbaik [15]. Untuk tahapan lebih jelasnya dapat dilihat pada gambar 1.

2.1 Gagal Jantung

Cardiovascular Disease (CVD) adalah penyakit yang disebabkan oleh gangguan pada kerja jantung dan pembuluh darah [16]. Menurut *New York Heart Association* (NYHA), gagal jantung dibedakan berdasarkan 4 kemampuan fisik. Derajat I pasien menunjukkan bisa beraktifitas secara normal, derajat II pasien menunjukkan gejala ringan saat melakukan aktivitas, derajat III pasien sudah mulai menunjukkan adanya keterbatasan fisik, dan derajat IV pasien sudah tidak bisa melakukan aktivitas apapun tanpa keluhan [17].

2.2 Data Mining

Data Mining merupakan inti dari proses *Knowledge Discovery in Database* (KDD). Dimana menggunakan data dengan melibatkan algoritma, mengembangkan model, dan menemukan pola yang sebelumnya tidak diketahui [18]. Nantinya data akan disimpan di database dan akan diproses secara otomatis dengan aplikasi menggunakan teknik dan perhitungan algoritma [19].



Gambar 1. Metodologi Penelitian

2.3 C4.5

Algoritma C4.5 juga dikenal sebagai decision tree dimana merupakan salah satu algoritma yang sering digunakan dalam teknik klasifikasi data mining [20]. Algoritma C4.5 biasanya akan melakukan prediksi dari berbagai informasi berdasarkan data yang digunakan menghitung berdasarkan atribut yang digunakan [21]. Nilai gain dari tiap atribut inilah yang menentukan node dalam pohon keputusan [20].

$$\text{Entropy (S)} = \sum_{i=1}^n - p_i * \log_2 p_i \quad (1)$$

Keterangan:

- S : Himpunan kasus
- N : Jumlah partisi S
- Pi : Proporsi dari Si dan S

$$\text{Gain(S, A)} = \text{Entropy} - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy (S}_i) \quad (2)$$

Keterangan:

- S : Himpunan kasus
- A : Atribut
- N : Jumlah partisi atribut A
- |Si| : Jumlah kasus pada partisi ke-i
- |S| : Jumlah kasus dalam S

2.4 Naïve Bayes Classifier (NBC)

Naïve Bayes Classifier merupakan salah satu metode klasifikasi yang menggunakan teori probabilitas untuk memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. [22]. Naïve bayes merupakan salah satu Algoritma yang banyak digunakan berdasarkan beberapa sifatnya yang sederhana, Algoritma ini mengklasifikasikan data berdasarkan probabilitas P atribut x dari setiap kelas y data [23]. Semua atribut pada setiap kategori metode Naïve Bayes dianggap bebas dan tidak memiliki ketergantungan satu sama lain [14].

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (3)$$

Keterangan:

- X : Data dengan kelas yang belum diketahui
- H : Hipotesis data X merupakan suatu kelas spesifik
- P(H | X) : Probabilitas hipotesis H berdasarkan kondisi X
- P(H) : Probabilitas hipotesis H

$P(X | H)$: Probabilitas X berdasarkan kondisi pada hipotesis H
 $P(X)$: Probabilitas hipotesis X

2.5 Support Vector Machine (SVM)

Support Vector Machine merupakan metode yang dapat digunakan untuk *klasifikasi linier*, metode SVM mengalami banyak perkembangan sehingga metode ini dapat melakukan klasifikasi yang sifatnya *non-linier*, konsep yang digunakan yaitu konsep karnel pada ruang kerja[24]. Pada hakikatnya SVM dilakukan dengan tujuan mendapatkan *hyperplane* yang mana himpunan data dipisahkan kepada dua kelas secara linier. Metode ini berusaha mendapatkan *hyperplane* yang paling optimal.[25]

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan dan Preprocessing Data

Sebelum data diproses dilakukan tahap *preprocessing* terdahulu pada data mining. *Preprocessing* data adalah proses yang dilakukan untuk memulai data mentah yang akan disiapkan untuk pemrosesan tahap selanjutnya [26]. pada penelitian ini data yang digunakan yaitu data prediksi gagal jantung yang bersumber dari webside Kaggle Dataset dengan 13 atribut dan 299 data. Adapun 13 atribut yang akan digunakan yaitu: *age, anaemia, creatinine_phosphokinase, diabetes, ejection_fraction, high_blood_pressure, platelets, serum_creatinine, serum_sodium, sex, smoking, time, death_event*. Atribut label yang akan digunakan pada data prediksi gagal jantung yaitu atribut *death_event*. Data yang nantinya akan digunakan merupakan hasil dari cleaning data bisa dilihat pada tabel 1.

Tabel 1. Cleaning Data

No	Age	Anaemia	Creatinine_Phosphokinase	Diabetes	...	Time	Death_Event
1	75	Tidak	582	Tidak	...	4	Ya
2	55	Tidak	7861	Tidak	...	6	Ya
3	65	Tidak	146	Tidak	...	7	Ya
4	50	Ya	111	Tidak	...	7	Ya
5	65	Ya	160	Ya	...	8	Ya
...
299	50	Tidak	196	Tidak	...	285	Tidak

Setelah dilakukannya Cleaning data, selanjutnya melakukan transformasi data. Data yang ada ditransformasi kedalam bentuk angka atau simbol yang dilakukan dengan tujuan mempermudah perhitungan, dari ke 13 atribut terdapat 6 atribut yang akan melalui tahap transformasi yaitu *high_blood_pressure, anemia, sex, death_event* dan *smoking*. Perubahan data dapat dilihat pada tabel 2.

Tabel 2. Data Transformasi

No	anaemia	diabetes	high_blood_pressure	sex	smoking	death_event
1	0	0	1	1	0	1
2	0	0	0	1	0	1
3	0	0	0	1	1	1
4	1	0	0	1	0	1
5	1	1	0	0	0	1
...
299	0	0	0	1	1	0

Berdasarkan atribut pada tabel 2 maka dapat diperoleh transformasi secara keseluruhan, yang ditunjukkan pada tabel 3:

Tabel 3. Transformasi Keseluruhan

No	Age	Anaemia	...	Diabetes	Ejection_Fraction	Death_Event
1	75	0	...	0	20	1
2	55	0	...	0	38	1
3	65	0	...	0	20	1
4	50	1	...	0	20	1
5	65	1	...	1	20	1
...
299	50	0	...	0	45	0

3.2 Klasifikasi Algoritma C4.5, SVM dan NBC

Pada penerapan ini dataset telah ditransformasi dan data telah dipisah menggunakan teknik pembagian data yaitu holdout, dimana perbandingan data training dan data testing yaitu 80:20, dari 299 data diperoleh 239 data *training* dan 60 data *testing*. *Google Colabs* merupakan *tools* yang akan digunakan untuk

implementasi ketiga algoritma tersebut. Terdapat dua kelas pada dataset gagal jantung yaitu “Ya” Meninggal dan “Tidak” Meninggal.

Selanjutnya melakukan perhitungan performa klasifikasi setiap algoritma. Hasil dari pengujian dinyatakan dalam bentuk nilai *recall*, *precision* dan *accuracy* yang didapat dari setiap algoritma kedalam bentuk confusion matrix yang dapat dilihat pada tabel 4 sampai 6.

Tabel 4. *Confusion Matrix C4.5*

	True Ya	True Tidak
Pred. Ya	37	9
Pred. Tidak	3	11

Confusion Matrix yang terdapat pada tabel 4. Menunjukkan bahwa informasi yang diberikan confusion matrix yaitu nilai *true positive*, dan nilai *true negative*. Dimana Prediksi Ya dengan *True positive* sebesar 37 dan 9 *True Negative*, sedangkan pada Prediksi Tidak dengan *True Positive* 3 data dan sebanyak 11 data pada *True Negative*.

Tabel 5. *Confusion Matrix SVM*

	True Ya	True Tidak
Pred. Ya	61	0
Pred. Tidak	28	1

Dapat dilihat dari tabel 5. Dimana prediksi Ya dengan *true positif* sebesar 61 data dan *true negatif* sebesar 0 data, prediksi Tidak dengan *True Ya* sebanyak 28 dan *True tidak* sebanyak 1 data.

Tabel 6. *Confusion Matrix NBC*

	True Ya	True Tidak
Pred. Ya	33	2
Pred. Tidak	13	12

Pada tabel *Confusion Matrix NBC*, diketahui bahwa pediksi Ya dengan *True Positif* sebanyak 33 dan *true Negatifnya* terdapat 2 data, sedangkan pada prediksi Tidak Dengan *True Ya* sebanyak 13 dan *True Negatif* sebanyak 12.

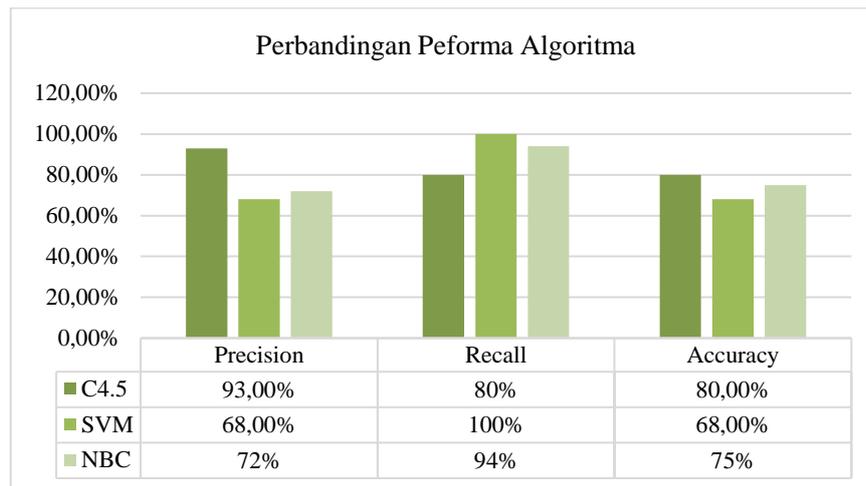
3.3 Perbandingan *Performance Algoritma*

Hasil pengujian *Confusion Matrix* dari kedua kelas label yang dapat dianalisis dari hasil rata-rata *recall*, *precision* dan *accuracy* dapat dilihat pada tabel 7.

Tabel 7. Hasil dan pengujian Setiap Algoritma

Metode	<i>Precision</i>	<i>Recall</i>	<i>Accuracy</i>
C4.5	0.93	0.8	0.80
SVM	0.68	100	0.68
NBC	0.72	0.94	0.75

Berdasarkan dari tabel 7 diketahui bahwa hasil klasifikasi menggunakan Google Colab dengan algoritma C4.5 diperoleh nilai akurasi algoritma C4.5 sebesar 80%, algoritma NBC sebesar 75% sedangkan algoritma SVM sebesar 68% Perbandingan *accuracy* algoritma C4.5, NBC dan SVM dapat dilihat pada gambar 2.



Gambar 2. Perbandingan Peforma Algoritma

Dari gambar 1. Berdasarkan pembahasan dan hasil penelitian didapat algoritma C4.5 merupakan algoritma yang dapat digunakan sebagai klasifikasi pada data tingkat kemungkinan orang meninggal diakibatkan penyakit gagal jantung. Berdasarkan hasil perhitungan, hasil pengujian menunjukkan algoritma C4.5 lebih optimal dibandingkan dengan algoritma SVM dan NBC untuk klasifikasi data pasien yang terkena gagal jantung dengan nilai accuracy sebesar 80%. Dari data testing didapat prediksi sebanyak 12 data beresiko kemungkinan meninggal dan sebanyak 48 data diprediksi tidak meninggal.

4. KESIMPULAN

Dari hasil yang sudah dilakukan dapat dikatakan bahwasannya penyakit gagal jantung merupakan penyakit dengan resiko kematian yang sangat besar. Oleh karena itu, perlu dilakukannya klasifikasi dimana disini peneliti menggunakan 3 algoritma yakni algoritma c4.5, SVM, dan NBC. Hasil klasifikasi yang dilakukan oleh tiga algoritma tersebut pada data penyakit gagal jantung dengan pembagian data menggunakan *holdout*. Memiliki nilai akurasi sebesar 80%, 68% dan 75%. Dapat diketahui nilai akurasi tertinggi didapat oleh algoritma C.45 sebesar 80%. Hal ini membuktikan bahwa algoritma C.45 lebih baik dibandingkan algoritma SVM dan NBC dalam memprediksi penyakit gagal jantung.

REFERENCES

- [1] Hernawati Basir, R. A. Rusmin, Yusriyani, Z. Thahir, and A. T. D. Pine, "Jurnal Pengabdian Kefarmasian Volume I, No.2, November 2020," *Pengabdi. Kefarmasian*, vol. 2, no. 2, pp. 32–37, 2021.
- [2] D. W. Adnyani and M. G. Juniarta, "Efektivitas Latihan Yoga Dalam Mengatasi Penyakit Jantung Koroner (Pjk)," *J. Yoga Dan Kesehat.*, vol. 3, no. 2, p. 129, 2020, doi: 10.25078/jyk.v3i2.1695.
- [3] D. Cahya Putri Buani, "Penerapan Algoritma Naïve Bayes dengan Seleksi Fitur Algoritma Genetika Untuk Prediksi Gagal Jantung," *EVOLUSI J. Sains dan Manaj.*, vol. 9, no. 2, pp. 43–48, 2021, doi: 10.31294/evolusi.v9i2.11141.
- [4] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [5] Adinda Maharani, Siti Fatimah Muis, Yushila Meyrina, Etisa Adi Murbawani, and Enny Porbasari, "Kesesuaian Antara A.S.P.E.N Malnutrition Guideline (A.S.P.E.N MG) dengan Subjective Global Assesment (SGA) pada Pasien Gagal Jantung Kronik Nyha 3-4," *JNH (Journal Nutr. Heal.* , vol. 8, no. 2, pp. 86–99, 2020.
- [6] D. Prihatiningsih and T. Sudyasih, "Perawatan Diri Pada Pasien Gagal Jantung," *J. Pendidik. Keperawatan Indones.*, vol. 4, no. 2, 2018, doi: 10.17509/jpki.v4i2.13443.
- [7] Harbanu H Mariyono and A. Santoso, "No GAGAL JANTUNG *Harbanu," *วารสารวิชาการมหาวิทยาลัยอีสเทิร์นเอเซีย*, vol. 4, no. 1, pp. 88–100, 2557.
- [8] D. Maulana and Rezayadi Yahya, "IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK KLASIFIKASI PENDERITA PENYAKIT JANTUNG DI INDONESIA MENGGUNAKAN RAPID MINER," *Биохимия*, vol. 84, no. 10, pp. 1511–1518, 2019, doi: 10.1134/s0320972519100129.
- [9] M. A. Bianto, K. Kusrini, and S. Sudarmawan, "Perancangan Sistem Klasifikasi Penyakit Jantung Menggunakan Naïve Bayes," *Creat. Inf. Technol. J.*, vol. 6, no. 1, p. 75, 2020, doi: 10.24076/citec.2019v6i1.231.
- [10] A. Sefharni, I. E. Hendrawan, and C. Rozikin, "KLASIFIKASI PENYAKIT JANTUNG DENGAN

- MENGGUNAKAN ALGORITMA C4.5,” vol. 7, no. 2, 2022.
- [11] N. G. Ramadhan and A. Khoirunnisa, “Klasifikasi Data Malaria Menggunakan Metode Support Vector Machine,” *J. Media Inform. Budidarma*, vol. 5, no. 4, p. 1580, 2021, doi: 10.30865/mib.v5i4.3347.
- [12] M. Irfan, N. Lukman, A. A. Alfauzi, and J. Jumadi, “Comparison of algorithm Support Vector Machine and C4.5 for identification of pests and diseases in chili plants,” *J. Phys. Conf. Ser.*, vol. 1402, no. 6, 2019, doi: 10.1088/1742-6596/1402/6/066104.
- [13] A. Qurotul, E. Tasia, N. Nazira, P. F. Pratama, M. R. Anugrah, and J. Adhiva, “Implementasi Algoritma Naïve Bayes Classifier (NBC) untuk Klasifikasi Penyakit Ginjal Kronik,” vol. 4, no. September, pp. 72–76, 2022, doi: 10.30865/json.v4i1.4781.
- [14] A. Desiani, M. Akbar, Irmeilyana, and A. Amran, “Implementasi Algoritma Naïve Bayes dan Support Vector Machine (SVM) Pada Klasifikasi Penyakit Kardiovaskular,” *J. Tek. Elektro dan Komputasi*, vol. 4, no. 2, pp. 207–214, 2022.
- [15] N. T. Luchia, H. Handayani, F. S. Hamdi, D. Erlangga, and S. Fitri Octavia, “Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 2, no. 2, pp. 35–41, 2022.
- [16] E. Nurlia and U. Enri, “Penerapan Fitur Seleksi Forward Selection Untuk Menentukan Kematian Akibat Gagal Jantung Menggunakan Algoritma C4.5,” *J. Tek. Inform. Musirawas) Elin Nurlia*, vol. 6, no. 1, p. 42, 2021.
- [17] H. Nursita and A. Pratiwi, “Peningkatan Kualitas Hidup Pada Pasien Gagal Jantung: A Narrative Review Article,” *J. Ber. Ilmu Keperawatan*, vol. 13, no. 1, p. 11, 2020, [Online]. Available: <https://doi.org/10.23917/bik.v13i1.11916>
- [18] A. Ridwan and T. N. Sari, “The comparison of accuracy between naïve bayes clasifier and c4.5 algorithm in classifying toddler nutrition status based on anthropometry index,” *J. Phys. Conf. Ser.*, vol. 1764, no. 1, 2021, doi: 10.1088/1742-6596/1764/1/012047.
- [19] M. E. Sulaiman *et al.*, “Analisis Penyebab Banjir di Kota Samarinda,” *J. Geogr. Gea*, vol. 20, no. 1, pp. 39–43, 2020, doi: 10.17509/gea.v20i1.22021.
- [20] H. I. Islam, M. K. Mulyadien, and U. Enri, “Penerapan Algoritma C4.5 dalam Klasifikasi Status Gizi Balita,” *Ilm. Wahana Pendidik.*, vol. 8, no. 10, pp. 116–125, 2022.
- [21] N. Sunanto and G. Falah, “Penerapan Algoritma C4.5 Untuk Membuat Model Prediksi Pasien Yang Mengidap Penyakit Diabetes,” *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 2, pp. 208–216, 2022, doi: 10.36341/rabit.v7i2.2435.
- [22] M. R. Yuliansyah, M. B, and A. Franz, “Perbandingan Metode K-Nearest Neighbors dan Naïve Bayes Classifier Pada Klasifikasi Status Gizi Balita di Puskesmas Muara Jawa Kota Samarinda,” *Adopsi Teknol. dan Sist. Informasi(ATASI)*, vol. 1, no. 1, pp. 8–20, 2022, [Online]. Available: <https://e-journals2.unmul.ac.id/index.php/atasi/article/view/25/30>
- [23] J. Homepage, P. Algoritma Klasifikasi untuk Analisis Sentimen, E. Ditendra, S. Romelah, M. Habil Arsyiddik Tanjung, and M. Sarah, “Perbandingan Algoritma Klasifikasi untuk Analisis Sentimen Islam Nusantara Indonesia,” vol. 2, no. April, pp. 71–77, 2022.
- [24] E. Ramon, A. Nazir, N. Novriyanto, Y. Yusra, and L. Oktavia, “Klasifikasi Status Gizi Bayi Posyandu Kecamatan Bangun Purba Menggunakan Algoritma Support Vector Machine (Svm),” *J. Sist. Inf. dan Inform.*, vol. 5, no. 2, pp. 143–150, 2022, doi: 10.47080/simika.v5i2.2185.
- [25] C. Chazar and B. Erawan, “Machine Learning Diagnosis Kanker Payudara Menggunakan Algoritma Support Vector Machine,” *Inf. (Jurnal Inform. dan Sist. Informasi)*, vol. 12, no. 1, pp. 67–80, 2020, doi: 10.37424/informasi.v12i1.48.
- [26] F. Akbar, H. Wira Saputra, A. Karel Maulaya, M. Fikri Hidayat, and Rahmaddeni, “Implementasi Algoritma Decision Tree C4.5 dan Support Vector Regression untuk Prediksi Penyakit Stroke,” vol. 2, no. October, pp. 61–67, 2022.