

Klasifikasi Penyakit Jantung Menggunakan Decision Tree dan KNN Menggunakan Ekstraksi Fitur PCA

Dewi Nasien^{1*}, Sirvan², Deny³, Ryan Syahputra⁴, Alberta Akbar Marunduri⁵, Richardo Prawinata See⁶,
¹⁻⁶ Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Institut Bisnis dan Teknologi Pelita Indonesia
e-mail: ¹[*1dewinasien@lecturer.pelitaindonesia.ac.id](mailto:dewinasien@lecturer.pelitaindonesia.ac.id), ²sirvan@student.pelitaindonesia.ac.id,
³deny@student.pelitaindonesia.ac.id, ⁴ryansyahputra@student.pelitaindonesia.ac.id,
⁵albertaakbarmarunduri@student.pelitaindonesia.ac.id,
⁶richardopravinatasee@student.pelitaindonesia.ac.id

Abstract – Heart disease, which is the main cause of death, is the focus of BPJS Health's treatment and financing. For preventive efforts, predicting heart disease in patients is an important step. In this research, the classification process was carried out using two methods, namely decision tree and KNN, to predict heart disease. The decision tree and KNN methods are approaches commonly used in the classification of heart disease. Decision trees build tree-based decision models, while KNN combines several decision trees to improve prediction performance and stability. The performance evaluation results of both methods can provide a comprehensive view of the effectiveness of each in predicting heart disease on the dataset used. Evaluation metrics such as accuracy, precision, recall, and F1 score will provide information about the extent to which the model is able to classify data correctly and identify heart disease cases well.

Keywords: K-fold Cross Validation, Clasification, Heart disease

Abstrak – Penyakit jantung, yang merupakan penyebab utama kematian, menjadi fokus penanganan dan pembiayaan BPJS Kesehatan. Untuk upaya preventif, prediksi penyakit jantung pada pasien menjadi langkah penting. Dalam penelitian ini, proses klasifikasi dilakukan menggunakan dua metode, yaitu decision tree dan KNN, untuk memprediksi penyakit jantung. Metode decision tree dan KNN merupakan pendekatan yang umum digunakan dalam klasifikasi penyakit jantung. Decision tree membangun model keputusan berbasis pohon, sedangkan KNN menggabungkan beberapa decision tree untuk meningkatkan kinerja dan kestabilan prediksi. Hasil evaluasi performa kedua metode dapat memberikan pandangan yang komprehensif tentang keefektifan masing-masing dalam memprediksi penyakit jantung pada dataset yang digunakan. Metrik evaluasi seperti akurasi, precision, recall, dan F1 score akan memberikan informasi tentang sejauh mana model mampu mengklasifikasikan data dengan benar dan mengidentifikasi kasus penyakit jantung dengan baik.

Kata Kunci – K-fold Cross Validation, Klasifikasi, Penyakit jantung.

I. PENDAHULUAN

Penyakit jantung adalah kondisi yang melibatkan gangguan pada organ jantung, yang dapat menunjukkan beberapa gejala yang perlu diwaspadai. Gejala-gejala tersebut mencakup rasa mual, muntah, keringat dingin, mudah lelah, sakit kepala, nyeri dada sebelah kiri, sesak napas, lemas, jantung berdebar, dan sensasi dada seperti diremas-remas [1], [2], [3], [4], [5]. Sebelum pandemi Covid-19, penyakit jantung telah menjadi penyebab utama kematian di dunia dan Indonesia. Data yang dirilis oleh WHO pada tahun 2021 mencatat 17.8 juta kematian akibat penyakit jantung, dengan 12.934.931 kasus terjadi pada tahun yang sama [6], [7], [8], [9] BPJS Kesehatan juga menjadi penyandang dana terbesar untuk penanganan penyakit jantung, dengan pembiayaan mencapai hampir 7.7 triliun rupiah. [3], [10], [11], [12], [13], [14] Mengingat angka kasus, kematian, dan biaya yang tinggi terkait dengan penyakit jantung, deteksi dini dan prediksi menggunakan klasifikasi machine learning menjadi suatu kebutuhan. Beberapa penelitian telah dilakukan dalam hal ini.

II. PENELITIAN YANG TERKAIT

Studi oleh Annisa (2019) menggunakan klasifikasi data mining untuk prediksi penyakit jantung dengan metode seperti decision tree, naïve bayes, k-nearest neighbors, KNN, dan decision stump. Hasil penelitian menunjukkan bahwa KNN memiliki akurasi tertinggi, mencapai 80.38% dengan 10-fold cross-validation dan uji-t [14], [15], [16], [17]

Penelitian oleh Azhima et al. (2022) tentang hybrid machine learning untuk memprediksi penyakit jantung menunjukkan bahwa model hybrid dapat meningkatkan akurasi. KNN mencapai akurasi 83.16%, dan dengan logistic regression mencapai 77.88%. Melalui model hybrid, akurasi klasifikasi mencapai 84.48%, menunjukkan peningkatan akurasi KNN sebesar 1.32% [18], [19], [20], [21]

Studi oleh Alham et al. (2021) mengenai sistem diagnosis penyakit jantung koroner dengan menggunakan decision tree di RSUD dr. Soedarso Pontianak mencapai nilai akurasi sebesar 94.4%. Penelitian ini membandingkan hasil klasifikasi decision tree dan KNN dengan k-fold cross-validation untuk menemukan prediksi terbaik terkait penyakit jantung [6], [22], [23][3].

III. METODE PENELITIAN

2.1 Sumber Data

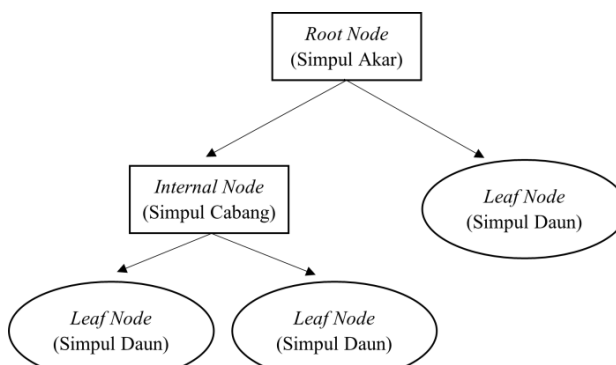
Dataset yang digunakan berjudul *Heart Disease* dan dipublikasikan di *Kaggle* [5]. Dataset terdiri dari 14 atribut yang ditunjukkan pada tabel 1. Kesimpulan terdapatnya penyakit jantung pada kolom *condition* dengan nilai 0 atau 1.

Tabel 1. Daftar Atribut Dataset

Atribut	Keterangan
Age	Umur
Sex	Jenis Kelamin
CP (Chest Pain)	Rasa sakit dada
Trestbps	Tekanan darah saat istirahat (mmHg)
Chol	Kolesterol (mg/dl)
Fbs	Gula darah puasa (>120mg/dl) (1=ya; 0=tidak)
Restecg	Hasil elektrokardiografi saat istirahat
Thalach	Detak jantung maksimal
Exang	Latihan yang diinduksi angina (1=ya;0=tidak)
Oldpeak	Depresi yang diinduksi oleh latihan relatif
Slope	Kemiringan puncak ST Segmen
Ca	Jumlah pembuluh darah yang berwarna setelah diwarnai flourosopy
Thal	Tipe kerusakan pembuluh darah (2=cacat sementara;1=cacat tetap;0=normal)
Condition	Indikasi penyakit jantung (1=ya;0=tidak)

2.2 Decision Tree

Decision tree adalah implementasi yang dirancang untuk menemukan keputusan dengan mempertimbangkan berbagai faktor yang terkait dengan masalah. Decision tree membangun pohon keputusan dengan menggunakan atribut sebagai akar, membuat cabang untuk setiap nilai kasus di cabang, dan mengulangi proses ini untuk setiap cabang hingga kasus di cabang memiliki kelas yang sama. Struktur decision tree dapat dilihat pada gambar 1.



Gambar 1. Struktur Decision Tree

Penentuan besarnya keefektifan suatu atribut dalam klasifikasi disebut *information gain* yang berdasarkan nilai *gain* tertinggi, menggunakan persamaan 3 [6].

$$Ga(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \dots \dots \dots (3)$$

dengan S adalah set kasus, A adalah atribut, N adalah sejumlah partisi atribut A , $|S_v|$ adalah jumlah kasus pada partisi ke- i , dan $|S|$ adalah jumlah kasus di S . Nilai *Entropy* dihitung sebelum nilai *gain* diperoleh. Rumus *Entropy* dituliskan pada persamaan 4.

$$Entro(S) = -p_{\oplus} \log_2 p_{\oplus} - p_{\ominus} \log_2 p_{\ominus} \dots\dots\dots (4)$$

2.3 K-Nearest Neighbor

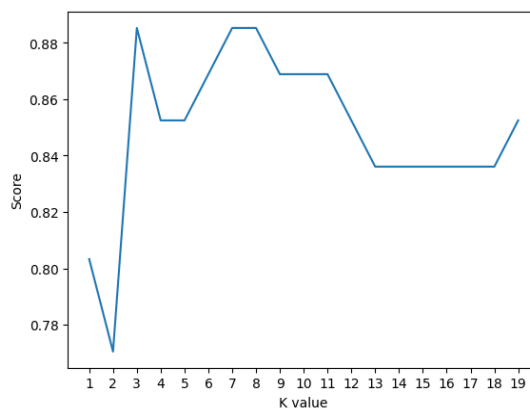
III.1 Algoritme K-Nearest Neighbor atau disingkat K-NN merupakan suatu algoritma non- parametric, supervised learning classifier yang menggunakan kedekatan untuk membuat klasifikasi atau prediksi tentang pengelompokan titik data individual[24].

$$X_{new} = \frac{X_{old} - X_{min}}{X_{max} - X_{min}}$$

Dimana:

X_{new} = Hasil normalisasi nilai data. X_{old} = Nilai data lama.

X_{min} = Nilai terkecil diantara kumpulan data dalam satu dimensi.



X_{max} = Nilai terbesar diantara kumpulan data dalam satu dimensi.

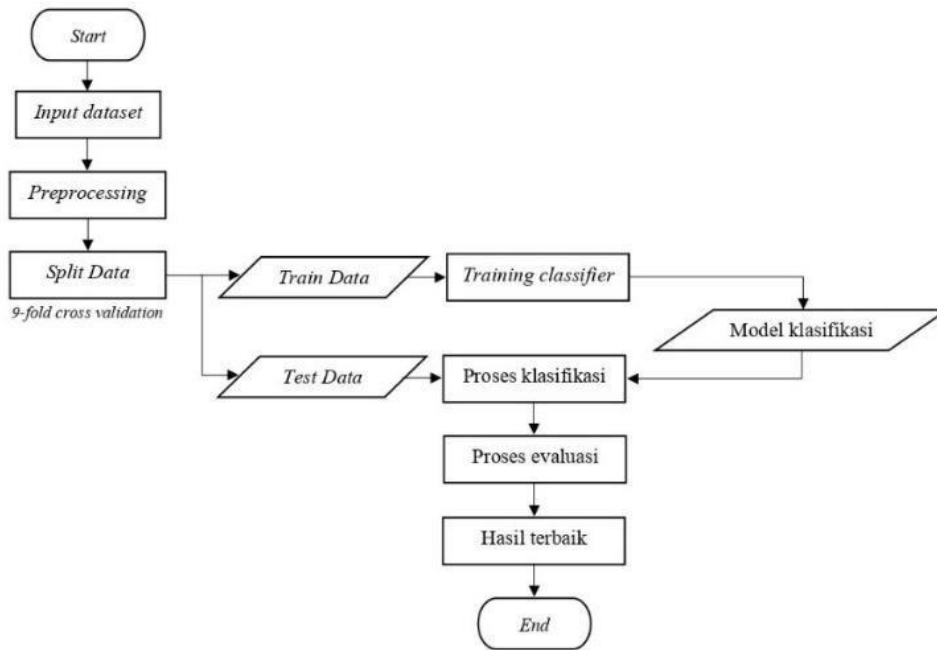
Gambar 2. Struktur KNN

2.4 K-Fold Cross Validation

K-fold cross validation adalah metode validasi pengujian sistematis yang melibatkan pembagian data menjadi k bagian, dan kemudian mengestimasi kesalahan pada setiap siklus pengujian. Langkah-langkah dari k-fold cross validation adalah sebagai berikut: Membagi jumlah data menjadi k bagian. Pada setiap siklus fold, salah satu bagian data menjadi data testing, sementara sisanya menjadi data training. Menghitung nilai evaluasi berdasarkan performa sistem pada bagian data testing tersebut. Melakukan langkah 2 dan 3 pada setiap fold, sehingga setiap bagian data menjadi data testing satu kali. Menghitung nilai evaluasi final dengan mengambil rata-rata dari nilai evaluasi pada setiap fold. Dengan demikian, k-fold cross validation memberikan gambaran yang lebih konsisten tentang performa sistem, karena melibatkan seluruh dataset sebagai bagian dari data testing dan training dalam beberapa iterasi pengujian.

2.5 Langkah Penelitian

Penelitian ini melibatkan beberapa tahap, antara lain, pembagian dataset menjadi k bagian sesuai dengan metode k-fold cross validation, melakukan preprocessing data, serta melakukan klasifikasi dengan menggunakan decision tree dan knn. Setelah itu, dilakukan perhitungan nilai evaluasi untuk memperoleh hasil klasifikasi terbaik. Alur keseluruhan dari penelitian ini dapat dilihat pada gambar 3.



Gambar 3. Alur Penelitian

IV. HASIL DAN PEMBAHASAN

3.1 Data

Dataset pada penelitian ini dilakukan pemeriksaan adanya *missing value* dengan menggunakan *softwareRstudio* didapatkan hasil pada gambar 4.

```

> colSums(is.na(heart))
  age      sex      cp      trestbps      chol      fbs      restecg      thalach      exang
  0         0         0         0         0         0         0         0         0
oldpeak      slope      ca      thal      condition
  0         0         0         0         0
  
```

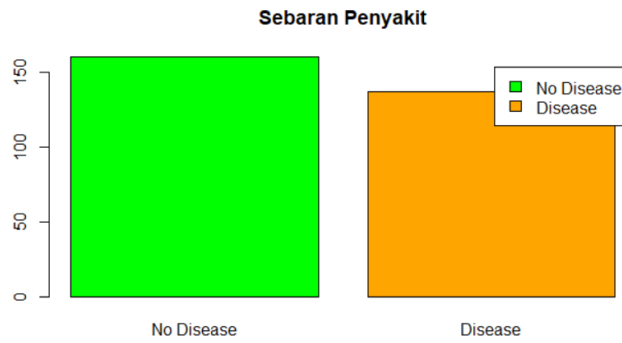
Gambar 4. Hasil Pemeriksaan *Missing Value*

Berdasarkan hasil yang diperoleh pada gambar 4, *dataset* tidak memiliki *missing value* sehingga dapat dilanjutkan ke tahap proses klasifikasi. Statistik deskriptif dari *dataset* yang digunakan ditunjukkan pada tabel 2.

Tabel 2. Statistik Deskriptif

	Min	1 st Qu	Median	Mean	3 rd Qu	Max
Age	29	48	56	54.54	61	77
Sex	0	0	1	0.6768	1	1
CP	0	2	2	2.158	3	3
Trestbps	94	120	130	131.7	140	200
Chol	126	211	243	247.4	276	564
Fbs	0	0	0	0.1448	0	1
Restecg	0	0	1	0.9966	2	2
Thalach	71	133	153	149.6	166	202
Exang	0	0	0	0.3266	1	1
Oldpeak	0	0	0.8	1.056	1.6	6.2
Slope	0	0	1	0.6027	1	2
Ca	0	0	0	0.6768	1	3
Thal	0	0	0	0.835	2	2
Condition	0	0	0	0.4613	1	1

Analisis data visualisasi dilakukan untuk mengetahui perbandingan rasio dari kelas. Gambar 5 menggambarkan *balanced data* dengan persentase pasien tidak mempunyai penyakit jantung sebesar 53.87% dan persentase pasien mempunyai penyakit jantung sebesar 46.13%.



Gambar 5. *Balanced Class Distribution*

Penelitian ini menggunakan *9-fold cross validation* dengan membagi *dataset* menjadi *9-fold* dengan ukuran yang sama, dimana *8-fold* akan digunakan sebagai data *training* dan *1-fold* akan digunakan sebagai data *testing*. Pembagian *dataset* menjadi data *training* dan data *testing* digambarkan pada tabel 3.

Tabel 3. *Pembagian Dataset*

<i>k-fold</i>	<i>Dataset</i>								
1	Test	Train	Train	Train	Train	Train	Train	Train	Train
2	Train	Test	Train	Train	Train	Train	Train	Train	Train
3	Train	Train	Test	Train	Train	Train	Train	Train	Train
4	Train	Train	Train	Test	Train	Train	Train	Train	Train
5	Train	Train	Train	Train	Test	Train	Train	Train	Train
6	Train	Train	Train	Train	Train	Test	Train	Train	Train
7	Train	Train	Train	Train	Train	Train	Test	Train	Train
8	Train	Train	Train	Train	Train	Train	Train	Test	Train
9	Train	Train	Train	Train	Train	Train	Train	Train	Test

3.2 Hasil Klasifikasi

Tahap pertama yang dilakukan untuk klasifikasi adalah melakukan pemodelan dengan data *training*. Setelah model klasifikasi pada data *training* terbentuk, model tersebut diuji menggunakan data *testing*. Performa klasifikasi dengan *decision tree* dapat dilihat pada tabel 4 dan performa klasifikasi dengan *KNN* dapat dilihat pada tabel 5.

Tabel 4. *Performa Klasifikasi Decision Tree*

	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>
<i>Validation 1</i>	72.73%	82.35%	70.00%	75.67%
<i>Validation 2</i>	81.82%	88.24%	78.95%	83.34%
<i>Validation 3</i>	75.76%	83.33%	75.00%	78.95%
<i>Validation 4</i>	84.85%	100.00%	73.68%	84.85%
<i>Validation 5</i>	87.88%	94.74%	85.71%	90.00%
<i>Validation 6</i>	81.82%	87.50%	77.78%	82.35%
<i>Validation 7</i>	66.67%	73.91%	77.27%	75.55%
<i>Validation 8</i>	72.73%	100.00%	65.38%	77.59%
<i>Validation 9</i>	72.73%	94.74%	69.23%	80.00%

Tabel 5. *Performa Klasifikasi KNN*

	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>
<i>Validation 1</i>	72.73%	82.35%	70.00%	75.67%
<i>Validation 2</i>	81.82%	88.24%	78.95%	83.34%
<i>Validation 3</i>	75.76%	83.33%	75.00%	78.95%
<i>Validation 4</i>	84.85%	100.00%	73.68%	84.85%
<i>Validation 5</i>	87.88%	94.74%	85.71%	90.00%
<i>Validation 6</i>	81.82%	87.50%	77.78%	82.35%
<i>Validation 7</i>	66.67%	73.91%	77.27%	75.55%
<i>Validation 8</i>	72.73%	100.00%	65.38%	77.59%
<i>Validation 9</i>	72.73%	94.74%	69.23%	80.00%

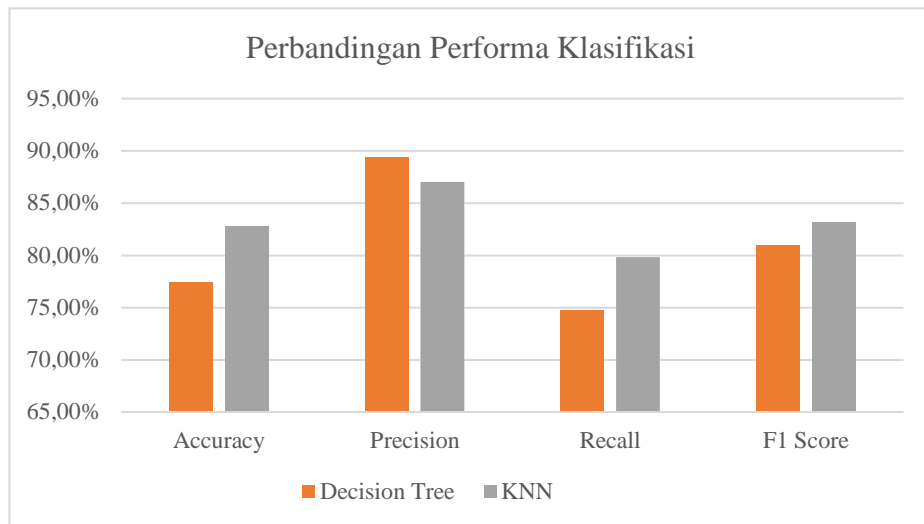
Klasifikasi *KNN* menunjukkan hasil *mtry* sebesar 2. *Mtry* menjelaskan jumlah variabel yang akan digunakan sebagai *split* pada setiap pohon yang terbentuk. *Ntree* pada klasifikasi ini adalah 500 yang merupakan nilai *default*

untuk *n*tree. *N*tree menunjukkan jumlah pohon yang dibentuk dalam klasifikasi *KNN*. Semakin banyak jumlah pohon yang dibentuk maka akan semakin baik. Berdasarkan perhitungan menggunakan *9-fold cross validation*, maka didapatkan performa klasifikasi penyakit jantung menggunakan *decision tree* dan *KNN* seperti yang dituliskan pada tabel 6.

Tabel 6. Performa Klasifikasi

	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>
<i>Decision Tree</i>	77.44%	89.42%	74.78%	80.92%
<i>KNN</i>	81.82%	87.04%	79.82%	83.13%

Berdasarkan pada tabel 6, perbandingan performa klasifikasi penyakit jantung menggunakan *decision tree* dan *KNN* juga digambarkan pada grafik pada gambar 6.



6. Perbandingan Nilai Evaluasi Klasifikasi

Berdasarkan gambar 6 diperoleh nilai akurasi *KNN* lebih tinggi dari nilai akurasi *decision tree*. *KNN* menghasilkan nilai akurasi sebesar 81.82% dan *decision tree* menghasilkan nilai akurasi sebesar 77.44%. Nilai *precision* pada *decision tree* adalah sebesar 89.42% dan nilai *precision* dari *KNN* adalah sebesar 87.04%. Nilai *recall* yang dihasilkan dari *KNN* adalah sebesar 79.82% dan nilai *recall decision tree* adalah sebesar 74.78%. Kemudian *F1 Score* dari *KNN* adalah sebesar 83.13% dan *F1 score decision tree* adalah sebesar 80.92%.

V. KESIMPULAN

Berdasarkan hasil analisis, disimpulkan bahwa klasifikasi penyakit jantung menggunakan metode *KNN* lebih baik dibandingkan dengan *decision tree*. Hal ini dikarenakan ketepatan klasifikasi dengan *KNN* memberikan hasil yang lebih baik. Secara khusus, klasifikasi dengan *KNN* menghasilkan nilai akurasi sebesar 81.82%, nilai *precision* sebesar 87.04%, nilai *recall* sebesar 79.82%, dan *F1 score* sebesar 83.13%. Dengan berbagai keterbatasan dalam penelitian ini maka beberapa saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai pada penelitian selanjutnya dapat menambah atau memperbarui metode atau algoritma yang digunakan sehingga hasil yang diperoleh dapat lebih akurat. Menggunakan *dataset* yang berjumlah lebih banyak untuk meningkatkan akurasi.

DAFTAR PUSTAKA

- [1] kemenkes, "Kementerian Kesehatan Republik Indonesia.," 2022.
- [2] R. Annisa, "Analisis Komparasi Algoritma Klasifikasi Data Mining untuk Prediksi Penderita Penyakit Jantung. Jurnal Teknik Informatika Kaputama (JTIK). No.1. Vol.3. 22-28.," 2019.
- [3] dan R. C. Alham S. R. J. I., Efy Y., "Sistem Diagnosis Penyakit Jantung Koroner Dengan Algoritma C4.5 Berbasis Website (Studi Kasus: RSUD Dr. Soedarso Pontianak). PETIR : Jurnal Pengkajian dan Penerapan Teknik Informatika. No.2. Vol.14. 214-222.," 2021.
- [4] Kaggle, "Kaggle. Heart Disease Cleveland UCI," 2022.
- [5] T. M. Mitchell, "Machine Learning. McGraw-Hill. Kementerian Kesehatan Republik Indonesia.," 2022.

- [6] I. P. Sari, D. Setiawan, and D. W. Marwan, "Sosialisasi Pelaksanaan Posyandu Melalui Transformasi Digital Kesehatan Menggunakan Aplikasi Mobile Posyandu Q," *J. Abdidas*, vol. 3, no. 5, pp. 870–876, 2022, doi: 10.31004/abdidas.v3i5.691.
- [7] D. Setiawan, R. N. Putri, and R. Suryanita, "Implementasi Algoritma Genetika Untuk Prediksi Penyakit Autoimun," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 4, no. 1, pp. 8–16, 2019, doi: 10.36341/rabit.v4i1.595.
- [8] I. Puspita Sari and D. Setiawan, "Aplikasi D'Laundry Berbasis Android Menggunakan Model Design Thinking Android-Based D'Laundry Application Using Model Design Thinking," *Juli*, vol. 2, no. 3, pp. 136–142, 2022.
- [9] D. Setiawan, D. Arisandi, and L. Trisnawati, "Aplikasi Prediksi Penyakit Sirosis Hati Menggunakan Algoritma Genetika," *J. SANTI (Sistem Inf. dan Teknol. Informasi)*, vol. 2, no. 1, pp. 31–40, 2022.
- [10] D. Azhima, "Hybrid Machine Learning Model untuk Memprediksi Penyakit Jantung dengan Metode Logistic Regression dan Random Forest. Jurnal Teknologi Terpadu. No.1. Vol.8. 40-46," 2022.
- [11] C. Y. Gobel, "Sistem Pakar Penyakit Liver Menggunakan K- Nearest Neighbors Algoritma," 2022.
- [12] N. T. Rahman, "Analisa Algoritma Decision Tree Dan Naïve Bayes Pada Pasien Penyakit Liver," *J. Fasilkom*, vol. 10, no. 2, pp. 144–151, 2020, doi: 10.37859/jf.v10i2.2087., 2022.
- [13] B. L. Trust, "Liver disease is now the biggest cause of death in those aged between 35-49 years old, new report reveals," British Liver Trust, 2019. [Online]. Available: <https://britishlivertrust.org.uk/liver-disease-is-now-the-biggest-cause-of-death-in-those-aged-between-35-49-years-old>, 2022.
- [14] and A. H. I. Setiawati, A. P. Wibowo, "Implementasi Decision Tree Untuk Mendiagnosis Penyakit Liver," *J. Inf. Syst. Manag.*, vol. 1, no. 1, pp. 13–17, 2019., 2019.
- [15] Rudiansyah, D. Setiawan, and Shucy Ramawati Fadila, "Identifikasi Faktor Risiko Stunting pada Anak-anak dengan Metode K means Clustering menggunakan Dataset Kaggle," *JEKIN - J. Tek. Inform.*, vol. 3, no. 1, pp. 10–16, 2023, doi: 10.58794/jekin.v3i1.443.
- [16] T. Yuhono, C. R. Hassolthine, and R. Sahara, "Prediksi Harga Steel Hot-Rolled (Hrc) Dengan Model Recurement Neural Network (Rnn)," vol. 4, no. 1, pp. 1–9, 2024.
- [17] T. Yuhono, C. R. Hassolthine, and R. Sahara, "PREDIKSI HARGA STEEL HOT-ROLLED (HRC) DENGAN MODEL RECUREMENT NEURAL NETWORK (RNN)," vol. 4, no. 1, pp. 1–9, 2024.
- [18] N. Maryam, D. Setiawan, and R. Noratama Putri, "Aplikasi Deteksi Dini Penyakit Tidak Menular Pada Lansia Dengan menggunakan metode Dynamic System Development Method (DSDM)," *JEKIN - J. Tek. Inform.*, vol. 3, no. 1, pp. 1–8, 2022, doi: 10.58794/jekin.v3i1.219.
- [19] Liza Trisnawati, D. Setiawan, and S. Ryanto, "ANTIDERM Dermatitis Consultation Application User Interface Design Model Using Design Sprint Method," *J. Int. Multidiscip.*, vol. 1, no. 1, pp. 1–17, 2023, doi: 10.58794/jim.v1i1.193.
- [20] D. Setiawan, R. N. Putri, and R. Suryanita, "Perbandingan Algoritma Genetika dan Backpropagation pada Aplikasi Prediksi Penyakit Autoimun," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 5, no. 1, pp. 21–27, 2019, doi: 10.23917/khif.v5i1.7173.
- [21] I. Engineering, D. Setiawan, M. C. Pratama, D. Arisandi, T. Informatika, and U. Abdurrah, "IMPLEMENTASI SISTEM KEAMANAN JARINGAN MENGGUNAKAN RULE-," vol. 7, no. 2, pp. 381–389, 2023.
- [22] P. Addie, D. Pengembangan, K. Pada, L. Elvitaria, D. Setiawan, and L. Susanti, "PENDAHULUAN Perkembangan ilmu pengetahuan dan teknologi yang sangat pesat saat ini menunjang dan mempermudah semua kegiatan di berbagai bidang dalam kehidupan manusia , salah satunya di Pendidikan . Dengan pemanfaatan teknologi media pembelajaran akan leb," vol. 8, no. 2, pp. 231–246, 2023.
- [23] M. P. Efendy and D. Setiawan, "PERANCANGAN APLIKASI MAKANAN EMPAT SEHAT LIMA SEMPURNA," vol. 5, no. 1, pp. 13–19, 2021.
- [24] Jack Billie Chandra and Dewi Nasien, "Application Of Machine Learning K-Nearest Neighbour Algorithm To Predict Diabetes," *Int. J. Electr. Energy Power Syst. Eng.*, vol. 6, no. 2, pp. 134–139, 2023, doi: 10.31258/ijeepse.6.1.117-121.