

ANALISIS DATA PENYAKIT JANTUNG MENGGUNAKAN ALGORITMA *DECISION TREE* DAN *SUPPORT VECTOR MACHINES* (SVM)

Fransiskus Dharma Hadi Wijaya

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara,
Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

E-mail: fransiskus.535220106@stu.untar.ac.id.

ABSTRAK

Penyakit jantung menjadi salah satu penyebab kematian utama seluruh dunia, sehingga analisis data menjadi semakin penting dalam upaya pencegahan dan pengobatan penyakit jantung. Dalam studi ini, analisis data penyakit jantung menggunakan dua algoritma *supervised learning* yaitu *Decision Tree* dan *Support Vector Machines* (SVM). Pada kesempatan kali ini, dataset yang digunakan adalah sebuah kumpulan data yang berisi berbagai variabel seperti usia, jenis kelamin, tekanan darah, hasil tes jantung dan yang lainnya. Langkah-langkah yang dilakukan meliputi pra-pemrosesan data, pemisahan dataset menjadi data pelatihan dan pengujian, pelatihan model, dan evaluasi performa model. Setelah dilakukannya perbandingan kinerja kedua metode dalam memprediksi keberadaan penyakit jantung pada pasien, hasil eksperimen menunjukkan bahwa kedua metode memberikan prediksi yang baik, namun SVM cenderung memiliki akurasi yang sedikit lebih tinggi dibandingkan *Decision Tree* pada dataset yang digunakan. Semoga hasil ini dapat bermanfaat serta menambah ilmu baru yang berharga bagi para medis yang profesional dan masyarakat umum dalam deteksi dini serta pengobatan penyakit jantung.

Kata kunci—*Decision Tree, Support Vector Machines, Machine Learning*

ABSTRACT

Heart disease is one of the main causes of death worldwide, so data analysis is becoming increasingly important in efforts to prevent and treat heart disease. In this study, heart disease data analysis uses two supervised learning algorithms, namely *Decision Trees* and *Support Vector Machines* (SVM). On this occasion, the dataset used was a collection of data containing various variables such as age, gender, blood pressure, heart test results and others. The steps taken include data pre-processing, separating the dataset into training and testing data, model training, and evaluating model performance. After comparing the performance of the two methods in predicting the presence of heart disease in patients, the experimental results show that both methods provide good predictions, but SVM tends to have slightly higher accuracy than *Decision Tree* on the dataset used. Hopefully these results can be useful and add valuable new knowledge for medical professionals and the general public in early detection and treatment of heart disease.

Keywords—*Decision Tree, Support Vector Machines, Machine Learning*

1. PENDAHULUAN

Penyakit jantung dan gangguan pembuluh darah terus menjadi ancaman global yang serius bagi kesehatan, menyebabkan tingkat kematian tertinggi di seluruh dunia. Berdasarkan data dari lembaga pusat statistik, sekitar 9,4 juta orang menjadi korban setiap tahunnya akibat penyakit kardiovaskular, dan diperkirakan jumlah ini akan terjadi kenaikan hingga 23,3 juta jiwa dio tahun 2030. Laporan dari Organisasi Kesehatan Dunia (WHO) pada tahun 2015 juga mencatat bahwa dari total 56,4 juta jiwa kematian yang terjadi, sekitar 39,5 juta disebabkan oleh penyakit yang tidak menular, di mana 17,7 juta di antaranya adalah akibat penyakit kardiovaskular. Sementara itu, kasus kematian yang terkait dengan penyakit jantung mencapai 7,4 juta setiap tahunnya, dengan perkiraan peningkatan menjadi 11 juta pada tahun mendatang. [1].

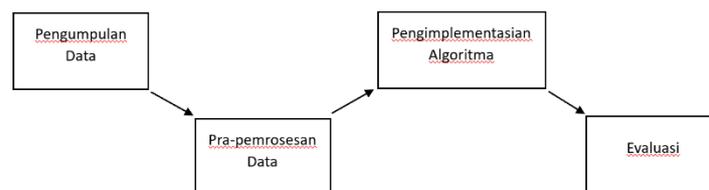
Setiap tahun, hampir 700.000 orang baru di Amerika masuk rumah sakit karena jantung koroner, dari laporan yang diterbitkan oleh American Heart Association (AHA). Dari jumlah tersebut, 40% mengalami kematian [2]. Penyakit kardiovaskular adalah penyakit yang sering menyebabkan korban jiwa seta cacat pada negara maju dan berkembang serta melibatkan gangguan yang terjadi pada jantung dan pembuluh darah [3].

Peran kesehatan masyarakat sangat berpengaruh dalam meningkatkan kesejahteraan masyarakat secara keseluruhan dan dalam upaya pencegahan penyakit jantung koroner. Terdapat sejumlah langkah yang dapat diambil untuk mengurangi kasus penyakit jantung koroner, termasuk melalui promosi kesehatan. Dalam konteks ini, promosi kesehatan dapat dilakukan dengan menyebarkan informasi tentang kesehatan jantung melalui berbagai saluran komunikasi, seperti poster atau platform media sosial, dengan tujuan meningkatkan kesadaran masyarakat tentang penyakit jantung koroner dan langkah-langkah pencegahannya. [4]. Berdasarkan data empiris, terdapat kecenderungan bahwa pria memiliki risiko terkena penyakit jantung koroner lebih tinggi hampir satu dekade lebih awal daripada wanita. Di sisi lain, pada wanita, risiko terkena penyakit jantung koroner meningkat ketika mereka memasuki masa menopause (berhentinya siklus menstruasi secara permanen) [5].

Penelitian ini dilakukan untuk menganalisis data terkait penyakit jantung yang bertujuan untuk memprediksi pola pola yang ada pada data tersebut dan untuk mengevaluasi tingkat keakuratan kinerja dari 2 algoritma *supervised learning Support Vector Machines (SVM)* dan *Decision Tree*.

2. METODE PENELITIAN

Dalam penelitian ini, tahapannya terbagi menjadi 4, yaitu pengumpulan data, pra-pemrosesan data, pengimplementasian algoritma klasifikasi, dan evaluasi. 2 algoritma yang digunakan yaitu *Decision Tree* dan *SVM*. Pada akhir analisis, setelah mendapatkan hasil klasifikasi menggunakan kedua algoritma tersebut, langkah selanjutnya adalah menguji validitasnya dengan menggunakan *confusion matrix* yang meliputi *precision*, *recall*, *f-1 score*, dan *accuracy* [6].



Gambar 1 Alur Penelitian

$$Akurasi = \frac{tp+tn}{tp+tn+fp+fn} \times 100\% \quad (1)$$

$$Precision = \frac{tp}{tp+fn} \times 100\% \quad (2)$$

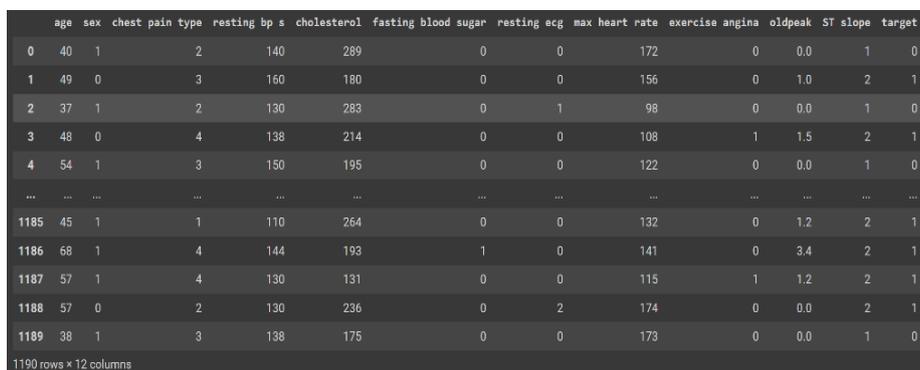
$$Recall = \frac{tp}{tp+fp} \times 100\% \quad (3)$$

$$F-1 \text{ Score} = \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Dalam persamaan yang telah disebutkan, tp menggambarkan jumlah prediksi yang benar positif, tn mengacu pada jumlah prediksi yang benar negatif, fp merujuk pada jumlah prediksi yang salah positif, dan fn melambangkan jumlah prediksi yang salah negatif [7].

2.1 Pengumpulan Data

Dataset Heart Deseas diambil dari *website Kaggle.com*. Data ini bertipe *integer* yang berisi 1190 baris dan 12 kolom yang mencakup informasi medis terkait penyakit jantung. Variabel-variabel yang termasuk dalam dataset ini mencakup usia, jenis kelamin, jenis nyeri dada (*Chest Pain Type*), tekanan darah sistolik (*Resting BP Systolic*), kadar kolesterol (*Cholesterol*), kadar gula darah (*Fasting Blood Sugar*), hasil elektrokardiogram (*Resting ECG*), detak jantung maksimum (*Max Heart Rate*), indikasi apakah pasien mengalami angina saat berolahraga (*Exercise Angina*), depresi segmen ST (*Oldpeak*), kemiringan segmen ST (*ST Slope*) dan Target.



	age	sex	chest pain type	resting bp s	cholesterol	fasting blood sugar	resting ecg	max heart rate	exercise angina	oldpeak	ST slope	target
0	40	1	2	140	289	0	0	172	0	0.0	1	0
1	49	0	3	160	180	0	0	156	0	1.0	2	1
2	37	1	2	130	283	0	1	98	0	0.0	1	0
3	48	0	4	138	214	0	0	108	1	1.5	2	1
4	54	1	3	150	195	0	0	122	0	0.0	1	0
...
1185	45	1	1	110	264	0	0	132	0	1.2	2	1
1186	68	1	4	144	193	1	0	141	0	3.4	2	1
1187	57	1	4	130	131	0	0	115	1	1.2	2	1
1188	57	0	2	130	236	0	2	174	0	0.0	2	1
1189	38	1	3	138	175	0	0	173	0	0.0	1	0

Gambar 2 Dataset Heart Deseas.

2.2 Pra-pemrosesan Data

Langkah pertama dalam pra-pemrosesan data yaitu dengan melakukan pengecekan *missing value* pada *dataset heart deseas* tersebut. Pengecekan ini sangat penting dikarenakan nilai null dapat menyebabkan masalah saat menerapkan analisis data atau membangun model machine learning.

```
df.isna().sum().sort_values(ascending=False)
```

```
age          0
sex          0
chest pain type  0
resting bp s  0
cholesterol  0
fasting blood sugar  0
resting ecg  0
max heart rate  0
exercise angina  0
oldpeak     0
ST slope    0
target      0
dtype: int64
```

Gambar 3 Pengecekan Missing Value

Dari hasil pengecekan tersebut, didapati bahwa tidak ada *missing value* pada data *heart deseas* tersebut. Setelah proses pengecekan *missing value*, tahap terakhir adalah melakukan pemisahan data menjadi data latih, juga data uji untuk evaluasi model.

```
[8] X = df.drop('target', axis = 1)
     Y = df['target']
```

Gambar 4 Pemisahan data uji dan data latih

Tahapan ini dilakukan untuk mengukur performa model secara obyektif dan memastikan bahwa model yang dibuat dapat diterapkan dengan baik pada data yang belum pernah terlihat.

2.3 Algoritma Support Vector Machine

SVM adalah algoritma yang bekerja dengan mencari *hyperplane* optimal. *Hyperplane* adalah fungsi yang dapat memisahkan dua kelas. *SVM* bertujuan untuk memaksimalkan *margin*, jarak antara sampel pelatihan dan batas Keputusan [8]. Vapnik memperkenalkan mesin vektor dukungan (*SVM*) tahun 1992 sebagai sebuah metode untuk klasifikasi efektif untuk menangani permasalahan nonlinier. *SVM* dianggap sebagai salah satu teknik pembelajaran mesin tercanggih setelah jaringan saraf (*NN*). [9].

Meskipun *SVM* pada awalnya didasarkan pada prinsip dasar klasifikasi linier, yaitu untuk kasus di mana kelas dapat dipisahkan secara linier, namun kemudian *SVM* telah diadaptasi untuk menangani masalah non-linier dengan memperkenalkan konsep kernel untuk mentransformasi data ke dalam ruang kerja berdimensi tinggi.

$$k(x_i, x_j) = x_i^T x_j \quad (5)$$

Persamaan (1) adalah rumus untuk kernel linier dalam konteks *Support Vector Machine (SVM)*. Dalam rumus ini, $k(x_i, x_j)$ adalah nilai kernel antara dua vektor fitur (x_i) dan (x_j), di mana (x_i) adalah vektor fitur dari sampel ke- i dan (x_j) adalah vektor fitur dari sampel ke- j . Operasi setelah samadengan (=) mengindikasikan perkalian titik antara kedua vektor fitur, yang menghasilkan nilai skalar yang mencerminkan kemiripan antara dua sampel dalam ruang fitur [10].

Dalam *SVM*, kernel linier seperti ini digunakan ketika data dapat dipisahkan dengan baik oleh *hyperplane* dalam ruang fitur asli. Kernel linier secara efisien mengukur kemiripan antara dua sampel dalam ruang fitur tanpa memerlukan pemetaan ke ruang fitur berdimensi tinggi. Setelah melakukan klasifikasi menggunakan *SVM*, langkah selanjutnya adalah menguji validitasnya menggunakan *confusion matrix*.

2.4 Algoritma Decision Tree

Decision Tree merupakan sebuah metode dalam menganalisis data. Metode ini biasa digunakan untuk mengklasifikasikan dan juga melakukan prediksi kepada variabel target dalam proses penambahan data. [11]. *Decision Tree* adalah metode pengambilan keputusan yang mengikuti struktur pohon hierarki yang terdiri dari akar sampai ke daun. Teknik ini juga digunakan untuk metode pembelajaran di bidang kecerdasan buatan yang dikenal dengan istilah *Decision Tree Learning (DTL)*. [12].

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{s_i}{S} * \text{Entropy}(S_i) \quad (6)$$

Dimana:

S : Sebuah himpunan suatu kasus

A : Atribut

n : Total partisi di dalam atribut A

| S_i | : Total kasus yang ada pada partisi ke- i

| S | : Total kasus yang ada di dalam S

Rumus *decision tree* yang disebutkan adalah rumus untuk menghitung gain dari suatu atribut. Gain adalah metrik yang digunakan untuk mengevaluasi atribut mana yang paling optimal untuk membagi data pada sebuah *decision tree* [13].

Pembentukan *Decision Tree* dapat dilakukan dengan cepat dibandingkan dengan metode klasifikasi lainnya. Pengklasifikasian menggunakan decision tree cenderung memberikan akurasi yang konsisten dan dalam beberapa kasus lebih baik daripada metode klasifikasi lainnya [14]. Sama seperti SVM, setelah melakukan klasifikasi menggunakan *Decision Tree*, selanjutnya adalah menguji validitasnya menggunakan *confusion matrix*.

2.5 Evaluasi

Pada tahap akhir ini, kita akan membandingkan hasil dari kinerja ke-2 algoritma dengan melihat algoritma mana yang mempunyai nilai terbaik dari seluruh percobaan yang dilakukan. Perbandingan dilakukan dengan melihat nilai rata – rata dari *precision*, *recall*, *F-1 Score*, dan *accuracy* pada setiap skenarionya.

3. HASIL DAN PEMBAHASAN

Penerapan algoritma SVM dan *Decision Tree* dilakukan menggunakan bahasa pemrograman Python. Program tersebut dilakukan sebanyak 15 kali percobaan model SVM dan 15 kali percobaan model *Decision Tree*. Percobaan tersebut dilakukan dengan 3 skenario. Skenario ke 1 dilakukan dengan data latih 70%, diikuti data uji 30%. Skenario ke 2 dilakukan dengan data latih 80%, diikuti data uji 20%. Skenario ke 3 dilakukan dengan data latih 90%, diikuti data uji 10%. Seluruh skenario menggunakan nilai $C = 0.15$. Untuk variabel target, dalam konteks prediksi penyakit jantung, angka 1 menunjukkan bahwa seseorang memiliki penyakit jantung dan angka 0 menunjukkan bahwa seseorang tidak memiliki penyakit jantung.

3.1 Hasil Implementasi Algoritma *Support Vector Machine* (SVM)

Pemrosesan data dilakukan menggunakan nilai $C = 0.15$ dan *random state* berbeda disetiap eksperimennya. Penelitian juga menggunakan 3 skenario berbeda dan berikut merupakan hasilnya.

Tabel 1 Hasil algoritma SVM skenario ke-1 dengan data latih 70% dan data uji 30%

Eksperimen	<i>P्रेसisi</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Akurasi</i>
1	0.83	0.83	0.83	0.83
2	0.84	0.84	0.84	0.84
3	0.84	0.83	0.84	0.84
4	0.85	0.85	0.85	0.85
5	0.81	0.80	0.80	0.80
Rata - Rata	0.83	0.83	0.83	0.83

Tabel 2 Hasil algoritma SVM skenario ke-1 dengan data latih 80% dan data uji 20%

Eksperimen	<i>P्रेसisi</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Akurasi</i>
1	0.82	0.82	0.82	0.82
2	0.85	0.85	0.85	0.85
3	0.83	0.83	0.83	0.83
4	0.86	0.86	0.86	0.86
5	0.83	0.83	0.83	0.83
Rata - Rata	0.84	0.84	0.84	0.84

Tabel 3 Hasil algoritma SVM skenario ke-1 dengan data latih 90% dan data uji 10%

Eksperimen	<i>Presisi</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Akurasi</i>
1	0.80	0.80	0.80	0.80
2	0.76	0.76	0.76	0.76
3	0.78	0.78	0.78	0.78
4	0.81	0.80	0.81	0.82
5	0.85	0.85	0.85	0.85
Rata - Rata	0.80	0.80	0.80	0.80

Dari hasil uji coba diatas, didapatkan bahwa nilai rata – rata dengan menggunakan algoritma SVM menyentuh angka 0.83 untuk tabel 1, 0.84 untuk tabel 2, dan 0.80 untuk tabel 3. Angka ini menunjukkan tingkat akurasi yang cukup baik dalam melakukan klasifikasi data.

3.2 Hasil Implementasi Algoritma Decision Tree

Pada algoritma *Decision Tree*, pemrosesan data juga dilakukan dengan menggunakan nilai $C = 0.15$ dan random state berbeda disetiap eksperimennya. Penelitian ini juga menggunakan 3 skenario berbeda dan berikut merupakan hasilnya.

Tabel 4 Hasil algoritma *Decision Tree* skenario ke-1 dengan data latih 70% dan data uji 30%

Eksperimen	<i>Presisi</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Akurasi</i>
1	0.82	0.82	0.82	0.82
2	0.76	0.76	0.76	0.76
3	0.78	0.78	0.78	0.78
4	0.83	0.83	0.83	0.83
5	0.80	0.79	0.79	0.80
Rata - Rata	0.80	0.80	0.80	0.80

Tabel 5 Hasil algoritma *Decision Tree* skenario ke-2 dengan data latih 80% dan data uji 20%

Eksperimen	<i>Presisi</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Akurasi</i>
1	0.87	0.84	0.84	0.84
2	0.84	0.83	0.84	0.84
3	0.83	0.82	0.83	0.83
4	0.85	0.84	0.84	0.84
5	0.80	0.79	0.79	0.79
Rata - Rata	0.84	0.82	0.83	0.83

Tabel 6 Hasil algoritma *Decision Tree* skenario ke-3 dengan data latih 90% dan data uji 10%

Eksperimen	<i>Presisi</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Akurasi</i>
1	0.75	0.75	0.75	0.76
2	0.71	0.71	0.71	0.71
3	0.80	0.80	0.80	0.80
4	0.81	0.81	0.81	0.82
5	0.83	0.82	0.82	0.82
Rata - Rata	0.80	0.80	0.80	0.80

Dari hasil uji coba diatas, didapatkan bahwa nilai rata – rata dengan menggunakan algoritma SVM menyentuh angka 0.80 untuk tabel 1, 0.82 – 0.84 untuk tabel 2, dan 0.80 untuk tabel 3. Angka ini menunjukkan tingkat akurasi yang cukup baik namun tidak setinggi algoritma SVM.

3.3 Hasil Analisis

Berdasarkan hasil analisis diatas, jika diuraikan satu persatu, pada skenario ke-1, algoritma SVM jauh lebih unggul dibandingkan *Decision Tree*. Hal ini ditunjukkan dengan hasil rata – rata dari precision, recall, f-1 score, dan accuracy yang menyentuh angka 0.83, 0.83, 0.83, 0.83 berturut turut.

Untuk skenario ke-2, algoritma SVM kembali unggul tipis dibandingkan dengan algoritma *Decision Tree*. Hasil skor rata – rata yang didapatkan dari algoritma SVM untuk *precision, recall, f-1 score, dan accuracy* adalah 8.4, 8.4, 8.4, 8.4 berturut turut. Sedangkan untuk algoritma *Decision Tree*, Hasil skor rata – rata yang didapatkan untuk *precision, recall, f-1 score, dan accuracy* adalah dikisaran 8.2 – 8.4.

Terakhir untuk skenario ke-3, algoritma SVM memperoleh skor rata – rata yang sama dengan algoritma *Decision tree*, yaitu untuk *precision, recall, f-1 score, dan accuracy* adalah 8.0, 8.0, 8.0, 8.0 berturut – turut.

4. KESIMPULAN

Kesimpulan dari penelitian ini dirangkum seperti berikut :

1. Algoritma Support Vector Machine (SVM) secara konsisten menunjukkan nilai *confusion matrix* yang lebih tinggi daripada algoritma *Decision Tree*. Ini menunjukkan bahwa SVM lebih efektif dalam memprediksi dan mengklasifikasikan data penyakit jantung dalam penelitian ini. Hal ini juga menandakan bahwa SVM mungkin merupakan pilihan yang lebih baik untuk analisis klasifikasi pada suatu *dataset*.
2. Penelitian berjalan dengan lancar, dengan performa yang lebih tinggi dalam berbagai metrik evaluasi, termasuk *convusion matrix*, SVM menunjukkan keandalan yang lebih tinggi dalam melakukan klasifikasi data penyakit jantung. Hal ini memberikan keyakinan lebih terhadap prediksi yang dihasilkan oleh model SVM dan menunjukkan potensi untuk penerapan yang lebih luas dalam praktek klinis atau penelitian medis.

Saran untuk pengembangan analisis data penyakit jantung adalah menjelajahi model klasifikasi alternatif selain *Decision Tree* dan SVM untuk melihat apakah ada model lain yang dapat memberikan performa yang lebih baik. Walaupun Support Vector Machine (SVM) [15].

DAFTAR PUSTAKA

- [1] L. Bachtiar, S. Maywati dan R. A. Gustaman, “FAKTOR RISIKO YANG BERHUBUNGAN DENGAN KEJADIAN PENYAKIT JANTUNG KORONER (PJK) (Analisis Data Sekunder di Rumah Sakit Umum Daerah Kabupaten Subang),” *Jurnal Kesehatan komunitas Indonesia*, vol. I, p. 53, 2023.
- [2] G. M. Ariaty, R. W. Sudjud dan R. H. Sitanggang, “Angka Mortalitas pada Pasien yang Menjalani Bedah Pintas Koroner berdasar Usia, Jenis Kelamin, Left Ventricular Ejection Fraction, Cross Clamp Time, Cardio Pulmonary Bypass Time, dan Penyakit Penyerta,” *Jurnal Anestesi Perioperatif*, p. 155, 2019.
- [3] L. F. Tampubolon, A. Ginting dan F. E. S. Turnip, “GAMBARAN FAKTOR YANG MEMPENGARUHI KEJADIAN PENYAKIT JANTUNG KORONER (PJK) DI PUSAT JANTUNG TERPADU (PJT),” *Jurnal Ilmiah Permas: Jurnal Ilmiah STIKES Kendal*, vol. 13, no. 3, p. 1044, 2023.
- [4] Melyani, L. N. Tambunan dan E. P. Barinbing, “Hubungan Usia dengan Kejadian Penyakit Jantung Koroner pada Pasien Rawat Jalan di RSUD dr. Doris Sylvanus Provinsi Kalimantan Tengah,” *Jurnal Surya Medika*, vol. 9, p. 120, 2023.

- [5] S. Retno, N. A. Wibowo dan I., "Identifikasi Faktor Kejadian Penyakit Jantung Koroner Terhadap Wanita Usia ≤ 50 Tahun di RSUD Haji Surabaya," *Jurnal Manajemen Asuhan Keperawatan*, vol. 6, p. 54, 2022.
- [6] D. G. Pradana, M. L. Alghifari, M. F. Juna dan S. D. Palaguna, "Klasifikasi Penyakit Jantung Menggunakan Metode Artificial Neural Network," *Indonesian Journal of Data and Science (IJODAS)*, vol. 3, no. 2, p. 56, 2022.
- [7] R. Hikmatulloh, H. M. P dan Q. Aini, "Penerapan Decision Tree untuk Prediksi Kepuasan Pengguna Bus Transjakarta," *Innovation in Research of Informatics (INNOVATICS)*, vol. 2, no. 2, p. 42, 2020.
- [8] A. Fatwa, "PENERAPAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) UNTUK KLASIFIKASI PENCEMARAN NAMA BAIK DI MEDIA SOSIAL TWITTER," *KOMPUTA: Jurnal Ilmiah Komputer dan Informatika*, vol. 12, no. 1, p. 75, 2023.
- [9] K. A. Rokhman, B. dan P. Arsi, "PERBANDINGAN METODE SUPPORT VECTOR MACHINE DAN DECISION TREE UNTUK ANALISIS SENTIMEN REVIEW KOMENTAR PADA APLIKASI TRANSPORTASI ONLINE," *JOISM: JURNAL OF INFORMATION SYSTEM MANAGEMENT*, vol. 2, no. 2, p. 3, 2021.
- [10] A. Kulsumarwati, I. Purnamasari dan B. A. Dermawan, "PENERAPAN SVM DAN INFORMATION GAIN PADA ANALISIS SENTIMEN PELAKSANAAN PILKADA SAAT PANDEMI," *Jurnal Teknologi Informatika dan Komputer MH. Thamrin*, vol. 7, no. 2, p. 104, 2021.
- [11] N. I. Nella, N. Y. Setiawan dan D. E. Ratnawati, "Klasifikasi Penerima Bantuan Program Keluarga Harapan menggunakan Algoritme Decision Tree C4.5 (Studi Kasus: Desa Mlirip Kabupaten Mojokerto)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 3, p. 1333, 2022.
- [12] S. P. Barus, "Penerapan Model Decision Tree pada Machine Learning untuk Memprediksi Calon Potensial Mahasiswa Baru," *Jurnal IKRAITH-INFORMATIKA*, vol. 6, no. 2, p. 60, 2022.
- [13] E. Fauziningrum dan E. I. Suryaningsih, "PENERAPAN DATA MINING METODE DECISION TREE UNTUK MENGUKUR PENGUASAAN BAHASA INGGRIS MARITIM (STUDI KASUS DI UNIVERSITAS MARITIM AMNI)," *Jurnal Saintek Maritim*, vol. 22, no. 1, p. 45, 2021.
- [14] B. H. Mawaridi dan M. Faisal, "Rekomendasi Merek Mobil Untuk Calon Pembeli Menggunakan Algoritma Decision Tree," *JURNAL INFORMATIKA*, vol. 10, no. 2, p. 158, 2023.
- [15] M. F. Asshiddiqi dan K. M. Lhaksamana, "Perbandingan Metode Decision Tree dan Support Vector Machine untuk Analisis Sentimen pada Instagram Mengenai Kinerja PSSI," *e-Proceeding of Engineering*, vol. 7, no. 3, p. 9942, 2020.