Komputika: Jurnal Sistem Komputer
Volume 13, Nomor 2, Oktober 2024, hlm. 149 - 156
DOI: **10.34010/komputika.v13i2.11330**

p-ISSN: 2252-9039
e-ISSN: 2655-3198

Klasifikasi Gagal Jantung Menggunakan Metode SVM (*Support Vector Machine*)

Laili Nur Farida[1*], Saiful Bahri[2]

[1)]Program Studi Matematika, Fakultas Sains Dan Teknologi, UIN Sunan Ampel Surabaya, Gn. Anyar, Surabaya, Indonesia 60294
[2)]Program Studi Biologi, Fakultas Sains Dan Teknologi, UIN Sunan Ampel Surabaya, Gn. Anyar, Surabaya, Indonesia 60294
**email: 09010220007@student.uinsby.ac.id*

(Naskah masuk: *21 Nopember 2023*; direvisi: *4 Juli 2024;* diterima untuk diterbitkan: *10 Oktober 2024*)

ABSTRAK – Gagal jantung adalah penyakit mematikan nomor satu di dunia. WHO dan WHF memperkirakan pada tahun 2025, penyakit jantung menjadi penyebab utama kematian di Asia. Saat ini, 78% kematian global akibat penyakit jantung terjadi pada orang miskin dan kelas menengah. Menurut RisKesDas KemenKes 2018, prevalensi gagal jantung di Indonesia mencapai 5%, lebih sering terjadi pada pria (66%) dibandingkan wanita (34%). Pasien gagal jantung di Indonesia yang meninggal saat perawatan rumah sakit mencapai 17.2%, dalam 1 tahun perawatan 11.3%, dan rehospitalisasi berulang 17%. Tujuan dari penelitian ini adalah melakukan klasifikasi penyakit gagal jantung menggunakan metode Support Vector Machine. Proses uji coba menghasilkan akurasi tertinggi pada kernel linear, RBF dan polynomial masing-masing sebesar 85.96%, 85.84%, dan 84.50%. Kernel yang menghasilkan akurasi paling tinggi, yaitu kernel linear dengan cost 0.1. Proses pengujian menggunakan parameter tersebut menghasilkan akurasi, presisi, recall, dan F1-score berturut-turut sebesar 89.13%, 86.21%, 96.15%, dan 90.91%. Berdasarkan hasil penelitian, diperoleh kesimpulan bahwa metode Support Vector Machine cukup baik dalam melakukan klasifikasi pada penyakit gagal jantung.

Kata Kunci – *Gagal Jantung, Klasifikasi, Penyakit, Supervised Learning, SVM*

Classification of Heart Failure using the SVM (Support Vector Machine) Method

ABSTRACT – *Heart failure is the number one killer disease in the world. WHO and WHF predict that by 2025, heart disease will be the leading cause of death in Asia. Currently, 78% of global deaths from heart disease occur in the poor and middle class. According to the Ministry of Health's 2018 RisKesDas, the prevalence of heart failure in Indonesia reached 5%, more common in men (66%) than women (34%). Heart failure patients in Indonesia who died during hospital treatment reached 17.2%, within 1 year of treatment 11.3%, and repeated rehospitalization 17%. The purpose of this study is to classify heart failure diseases using the Support Vector Machine method. The stages in this research consist of data conversion, outlier detection, normalization, data division, training process and testing process and evaluation of classification results. The test process produces the highest accuracy on linear, RBF and polynomial kernels of 85.96%, 85.84%, and 84.50%, respectively. The kernel that produces the highest accuracy is the linear kernel with cost 0.1. The testing process using these parameters resulted in accuracy, precision, recall, and F1-score of 89.13%, 86.21%, 96.15%, and 90.91%, respectively. Based on the results of the study, it is concluded that the Support Vector Machine method is quite good in classifying heart failure disease.*

Keywords – *Heart Failure, Classification, Disease, Supervised Learning, SVM*

1. PENDAHULUAN

Gagal jantung adalah keadaan serius ketika jantung tidak dapat memompa cukup darah untuk memenuhi kebutuhan tubuh [1]. Gagal jantung biasanya disebabkan oleh gangguan kesehatan yang berkaitan dengan organ jantung seperti gangguan irama jantung, penyakit jantung koroner dan lain-lain [2]. Gagal jantung merupakan penyakit mematikan nomor satu di dunia [3]. Penyakit ini dapat menyebabkan kematian jika tidak ditangani dengan baik. Gagal jantung menyebabkan rawat inap meningkat sekitar 6,5 juta per tahun [4]. Menurut data WHO (*World Health Organization*) dan WHF (*World Heart Federation*), pada tahun 2025 diperkirakan penyakit jantung akan menjadi penyebab utama kematian di negara-negara Asia. Tahun ini, setidaknya 78% angka kematian global disebabkan oleh penyakit jantung yang terjadi pada orang miskin dan kelas menengah. Angka kematian akibat penyakit jantung dari tahun 1990 hingga 2020 di negara berkembang akan meningkat sebesar 120% pada wanita dan 137% pada pria, sedangkan di negara maju peningkatan penyakit tersebut akan lebih rendah, yakni 29% pada wanita dan 48% pada pria [5]. Data RisKesDas (Riset Kesehatan Dasar) KemenKes (Kementerian Kesehatan) RI tahun 2018, prevalensi gagal jantung di Indonesia berdasarkan diagnosis dokter diperkirakan mencapai 5%, di mana lebih sering terjadi pada pria yaitu sebanyak 66% dibandingkan wanita yang hanya 34%. Pasien yang mengalami gagal jantung kemudian meninggal di Indonesia saat perawatan rumah sakit sebanyak 17.2%, sementara pasien yang meninggal dalam 1 tahun perawatan sebanyak 11.3% dan pasien yang mengalami rehospitalisasi berulang alias keluar-masuk rumah sakit sebanyak 17% [6].

Berdasarkan data WHO dan KemenKes RI terkait kasus penyakit gagal ginjal, diketahui bahwa penyakit tersebut sangatlah mematikan sebab tidak ditangani dengan baik. Oleh sebab itu, dibutuhkan suatu sistem untuk mengetahui ciri atau jenis penyakit gagal jantung agar dapat mencegah kematian. Salah satu cara untuk mengklasifikasi penyakit gagal jantung adalah menggunakan metode SVM (*Support Vector Machine*). Terdapat banyak penelitian yang menggunakan metode SVM yang bertujuan untuk klasifikasi. Salah satu penelitian terkait klasifikasi yang dilakukan oleh Vijaya dkk. terkait citra mammogram kanker payudara memperoleh hasil akurasi sebesar 94% [7]. Penelitian lain terkait klasifikasi opini film menggunakan SVM dan metode optimasi *Firefly* menghasilkan akurasi terbaik sebesar 89% [8].

Metode SVM merupakan model *supervised learning* yang awalnya dirancang untuk tugas klasifikasi biner dan memiliki kemampuan generalisasi berkualitas tinggi untuk masalah klasifikasi biner [9]. SVM mencoba memecahkan masalah klasifikasi dengan membentuk *hyperplane* yang memaksimalkan *margin* dengan membagi data ke dalam kelas-kelas. Jarak terdekat dari *hyperplane* ke titik masing-masing kelas dikenal sebagai *margin* [10]. Klasifikasi menggunakan SVM pada dataset yang berdimensi tinggi memiliki generalisasi yang baik, sehingga dapat memprediksi data baru yang tidak tedapat pada data pelatihan [11]. Teori matematika yang solid merupakan dasar dari SVM, seperti optimasi dan kernel sehingga kinerja dan justifikasi penggunaaan metode ini memiliki dasar yang kuat [12]. SVM juga dapat beradaptasi pada masalah yang *non-linear* dengan menerapkan trik kernel, seperti polynomial. Hal ini dapat membantu SVM untuk menentukan batas keputusan yang kompleks [13].

Dari beberapa metode penelitian yang ada, metode SVM mampu memperoleh hasil akurasi yang lebih tinggi dibandingkan dengan metode lain, seperti pada penelitian yang membandingkan algoritma SVM dengan KNN (*K-Nearest Neighbors*). Hasil klasifikasi pada penelitian tersebut menunjukkan bahwa metode SVM sangat baik dalam melakukan klasifikasi penyakit jantung dengan hasil akurasi tanpa normalisasi sebesar 84,61% dan dengan normalisasi sebesar 90,10%, sedangkan algoritma KNN menunjukkan akurasi tanpa normalisasi sebesar 64,83% dan dengan normalisasi sebesar 81,31% [14]. Penelitian lain yang membandingkan SVM dengan *Random Forest* terkait klasifikasi kanker payudara, SVM memperoleh hasil akurasi sebesar 95% dibandingkan dengan *Random Forest* yang menghasilkan akurasi 90% [10].

Berdasarkan penelitian-penelitian sebelumnya yang menggunakan SVM, diketahui bahwa metode tersebut cocok dalam melakukan klasifikasi. Cara kerja pada metode tersebut sangat cocok digunakan pada klasifikasi biner. Tujuan dari penelitian ini adalah untuk mengembangkan model klasifikasi penyakit gagal jantung yang lebih akurat menggunakan SVM. Selain itu, penelitian ini juga bertujuan untuk mengatasi masalah seperti mendeteksi *outlier*. Kontribusi dari penelitian ini adalah implementasi dan evaluasi model SVM dengan berbagai kernel dan pengujian model pada dataset yang berbeda untuk memastikan generalisasi model yang diusulkan dalam klasifikasi penyakit gagal

2. METODE DAN BAHAN

Data yang digunakan dalam penelitian ini yaitu dataset hasil laboratorium untuk penyakit gagal jantung yang diperoleh dari situs Kaggle (https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction) dengan format csv. Data

tersebut terdiri dari 12 atribut, 11 di antaranya sebagai atribut prediksi yaitu *Age*, *Sex*, *Chest Pain Type*, *Resting* BP, *Cholesterol*, *Fasting* BS, *Resting* ECG, *Max* HR, *Exercise Angina*, *Old peak* dan ST *Slope* serta satu atribut target yaitu *Heart Disease* yang terdiri dari 2 kelas yaitu 1 yang berarti penderita gagal jantung dan 0 yang berarti normal. Data yang tersedia berjumlah 918 data di mana 410 data masuk ke dalam kelas penderita gagal jantung dan 508 data masuk ke dalam kelas normal. Atribut data yang digunakan dalam penelitian ini dapat dilihat pada tabel 1.

Tabel 1. Atribut Data

Atribut	Keterangan	Variabel	Rentang Nilai
Age	Umur	Numerik	(28, 77]
Sex	Jenis kelamin	Kategori	F or M
Chest Pain Type	Nyeri dada	Kategori	ATA, ASY or NAP
Resting BP	Tekanan darah	Numerik	(0, 200]
Cholesterol	Kolesterol	Numerik	(0, 603]
Fasting BS	Gula darah puasa	Numerik	[0, 1]
Resting ECG	Elektrokardiogram	Kategori	ST, LVH or Normal
Max HR	Detak jantung maksimum	Numerik	(60, 202]
Exercise Angina	angina	Kategori	Y or N
Old peak	Depresi	Numerik Float	(-2.6, 6.2]
ST *Slope*	Kemiringan latihan	Kategori	Up, Down or Flat
Heart Disease	Kelas gagal jantung	Numerik	[0, 1]

Data penyakit gagal jantung memiliki data tipe karakter sehingga harus dilakukan *preprocessing* yang pertama yaitu berupa konversi data menjadi tipe numerik.

Setelah melakukan konversi, dilakukan deteksi *outlier* pada data terlebih dahulu menggunakan Persamaan 1 sampai Persamaan 3.

$$IQR_{min} = Q_1 - 1{,}5 \times IQR \quad (1)$$
$$IQR_{max} = Q_3 - 1{,}5 \times IQR \quad (2)$$
$$IQR = Q_3 - Q_1 \quad (3)$$

Dimana Q_1 merupakan nilai kuartil satu pada data dan Q_3 merupakan nilai kuartil tiga pada data. Suatu data dikatakan *outlier* jika data tersebut lebih kecil dari nilai minimum IQR atau lebih besar dari nilai maksimum IQR [15].

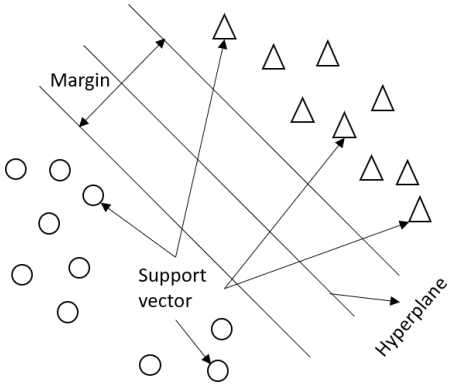
Preprocessing selanjutnya yaitu normalisasi agar dapat menangani masalah *outlier* dengan menggabungkan semua nilai *input* pada skala umum. Salah satu teknik normalisasi yaitu normalisasi *z-score* atau standarisasi. Normalisasi *z-score* ditunjukkan pada Persamaan 4.

$$x_{z-score} = \frac{x - mean(x)}{std_{dev}(x)} \quad (4)$$
$$std_{dev}(x) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x - mean(x))^2} \quad (5)$$

Dimana $x_{z-score}$ merupakan data hasil normalisasi, x merupakan data yang dinormalisasi, $mean\,(x)$ merupakan rata-rata dari x, $std_{dev}(x)$ merupakan standar deviasi dari x dan N merupakan jumlah data.

Proses selanjutnya adalah pelatihan dan pengujian menggunakan SVM berdasarkan *input* dan target pada Tabel 1. Proses pelatihan menggunakan data latih dan data uji dengan persentase 90:10.

SVM merupakan salah satu algoritma *deep learning* dengan model *supervised learning* yang berguna sebagai algoritma klasifikasi dan regresi [16]. Algoritma SVM termasuk metode klasifikasi linear dengan memaksimal *margin* dan menggunakan *hyperplane* dalam melakukan klasifikasi data secara biner [17]. Memaksimalkan nilai *margin* merupakan metode yang digunakan untuk mendapatkan nilai *hyperplane* yang optimal [18]. *Margin* merupakan jarak antara posisi terdekat di setiap kelas ke *hyperplane* [14]. *Hyperplane* merupakan batas antara dua kelas yang telah ditentukan [19]. Gambar *hyperplane* yang optimal dapat dilihat pada Gambar 1 [20].



Gambar 1. *Hyperplane* SVM

Pembentukan *hyperplane* dipengaruhi oleh pemilihan parameter *cost* (C) dan jenis kernel. *Cost* (C) merupakan parameter dalam SVM yang mengatur sejauh mana model harus memberikan penalti terhadap kesalahan klasifikasi pada data pelatihan [21]. Nilai C mengontrol *trade-off* antara *margin* yang lebar dan kesalahan klasifikasi pada

data pelatihan. Nilai C yang tinggi pada model bertujuan untuk meminimalkan kesalahan klasifikasi pada data pelatihan, sehingga menyebabkan *overfitting* karena model menjadi sangat sensitif terhadap data pelatihan. Nilai C yang rendah pada model bertujuan untuk memaksimalkan *margin decision boundary*, sehingga mengakibatkan *underfitting* karena model menjadi terlalu sederhana [22].

Kernel merupakan sebuah komponen yang bertanggung jawab untuk memetakan *input* berdimensi rendah ke ruang berdimensi lebih tinggi [9]. Kernel yang sering digunakan, antara lain kernel linear, kernel RBF (*Radial Basis Function*) dan kernel polinomial. Kernel linear merupakan fungsi kernel yang paling sederhana dan tidak memiliki parameter [23]. Fungsi kernel linier membuat sampel data dapat dipisahkan secara linier, kecil, dan durasi pelatihan yang relatif singkat [24]. Kernel linear dirumuskan pada Persamaan 6 [25].

$$K(x_i, x_j) = (x_i{}^T \times x_j) + c \quad (6)$$

Dimana $x_i{}^T$ merupakan *input* ke i yang di-*tranpose*, x_j merupakan *input* ke j dan c adalah parameter konstanta.

Kernel RBF adalah kernel yang paling populer di antara semua kernel di SVM [26]. Kernel RBF meliputi kernel gaussian, dimana kernel gaussian RBF bertujuan untuk mengukur kemiripan antar sampel [27]. Kernel gaussian dirumuskan pada Persamaan 7 [28].

$$K(x_i, x_j) = exp\left(-\frac{||x_i - x_j||^2}{2\sigma^2} + 1\right) \quad (7)$$

Dimana σ merupakan varian dan *hyperparameter* yang digunakan serta $||x_i - x_j||$ merupakan jarak *Euclidean* antara dua titik x_i dan x_j.

Kernel polinomial didefinisikan sebagai kernel non-stasioner. Kernel ini sangat cocok untuk kumpulan data pelatihan yang dinormalisasi [29]. Kernel polinomial digunakan ketika sampel lebih sedikit [30]. Parameter yang dapat diatur disebut dengan *slope alpha*, parameter konstanta dan derajat polinomial [31]. Kernel polinomial dirumuskan pada Persamaan 8 [32].

$$K(x_i, x_j) = [(\sigma x_i{}^T \times x_j) + c]^d \quad (8)$$

Dimana d merupakan derajat dari fungsi kernel polinomial.

Evaluasi model yang berkaitan dengan klasifikasi dapat menggunakan *Confusion Matrix*. *Confusion Matrix* merupakan bentuk penyajian hasil dengan matriks khusus [33]. Informasi pada kasus klasifikasi biner memuat TP (*True Positive*), TN (*True Negative*), FP (*False Positive*) dan FN (*False Negative*) [34]. TP merupakan data positif benar. TN merupakan data negatif benar. FP merupakan data negatif salah. FN merupakan data positif salah [35]. Gambar 2 menunjukkan kerangka *Confusion Matrix* [36].

Tabel 2. Kerangka *Confusion Matrix*

		Predicted Class	
		Positive	Negative
Actual Class	Positive	TP	FN
	Negative	FP	TN

Evaluasi model yang banyak digunakan adalah akurasi, presisi, *recall* dan *F1-score*.

Metrik yang paling sering digunakan dalam klasifikasi biner adalah akurasi. Akurasi merupakan metrik yang menunjukkan keakuratan suatu model pada klasifikasi biner yang telah dilakukan [34]. Kesimpulannya, akurasi merupakan persentase kedekatan nilai prediksi dengan nilai aktual [37]. Rumus akurasi ditunjukkan pada Persamaan 9.

$$Akurasi = \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\% \quad (9)$$

Presisi menunjukkan hasil keakuratan antara data aktual dengan data prediksi [38]. Presisi merupakan rasio data positif yang diprediksi benar dibandingkan dengan keseluruhan data yang diprediksi positif [39]. Rumus presisi ditunjukkan pada Persamaan 10.

$$Presisi = \frac{TP}{TP+FP} \times 100\% \quad (10)$$

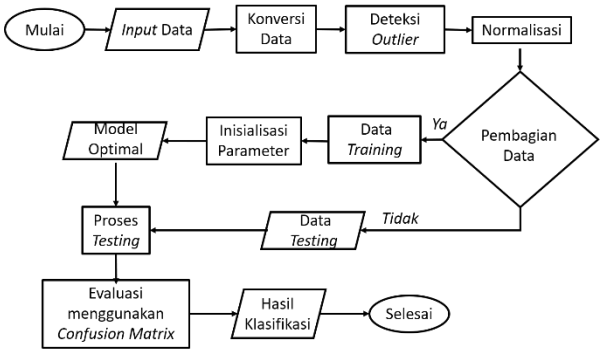
Recall menunjukkan tingkat keberhasilan model dalam memprediksi suatu data [40]. *Recall* merupakan rasio data positif yang diprediksi benar dengan keseluruhan data positif [41]. Rumus *recall* ditunjukkan pada Persamaan 11.

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (11)$$

F1-Score merupakan perbandingan rata-rata presisi dan *recall*, dimana perhitungan ini tidak bergantung pada jumlah data negatif yang diprediksi benar [34]. Rumus *F1-score* ditunjukkan pada Persamaan 12.

$$F1\ Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (12)$$

Gambar 3 menunjukkan diagram alir klasifikasi gagal jantung menggunakan SVM (*Support Vector Machine*).



Gambar 2. Diagram alir penelitian

3. HASIL DAN PEMBAHASAN

Data penyakit gagal jantung memiliki data tipe karakter sehingga harus dilakukan *preprocessing* yang pertama yaitu berupa konversi data menjadi tipe numerik. Setelah melakukan konversi, dilakukan deteksi *outlier* pada data terlebih dahulu menggunakan Persamaan 1 sampai Persamaan 3. Setelah melakukan deteksi *outlier*, *preprocessing* selanjutnya yaitu normalisasi menggunakan Persamaan 4 agar dapat menangani masalah *outlier* dengan menggabungkan semua nilai *input* pada skala umum. Tabel 2 menunjukkan hasil *preprocessing*.

Tabel 3. Hasil *Preprocessing*

Variabel	Data ke-1	Data ke-2	Data ke-3	...	Data ke-918
Age	40	49	37	...	38
Sex	1	0	1	...	1
ChestPain Type	1	2	1	...	2
Resting BP	0.41	1.49	-0.13	...	0.30
Cholesterol	0.82	-0.17	0.77	...	-0.22
Fasting BS	0	0	0	...	0
Resting ECG	1	1	2	...	1
Max HR	1.38	0.75	-1.52	...	1.42
Exercise Angina	0	0	0	...	0
Old peak	-0.83	0.11	-0.83	...	-0.83
ST *Slope*	2	1	2	...	2
Stage	0	1	0	...	0

Proses selanjutnya adalah pelatihan dan pengujian menggunakan SVM berdasarkan *input* dan target pada Tabel 2. Proses pelatihan menggunakan data latih dan data uji dengan persentase 80:20. Selain pembagian data latih dan data uji, parameter yang digunakan pada tahap pelatihan yaitu *cost* dan jenis kernel berdasarkan Persamaan 6 sampai Persamaan 8. Hasil pelatihan menghasilkan nilai presisi, F1-*score*, akurasi dan *recall* menggunakan Persamaan 9 sampai Persamaan 12, seperti pada Tabel 3 sampai Tabel 5.

Tabel 4. Hasil Uji Coba Kernel Linear

Kernel		Linear				
C		0.01	0.1	1	10	100
Hasil Evaluasi	Akurasi (%)	85.29	85.83	**86.92**	**86.92**	**86.92**
	Presisi (%)	86.90	86.49	**86.92**	**86.92**	**86.92**
	Recall (%)	86.03	87.78	**89.53**	**89.53**	**89.53**
	F1-Score (%)	86.47	87.13	**88.21**	**88.21**	**88.21**

Tabel 5. Hasil Uji Coba Kernel RBF

Kernel		RBF				
C		0.01	0.1	1	10	100
Hasil Evaluasi	Akurasi (%)	54.63	61.99	72.75	85.29	**86.24**
	Presisi (%)	54.63	60.66	72.09	86.17	**86.59**
	Recall (%)	100	86.53	81.80	87.03	**88.53**
	F1-Score (%)	70.66	71.33	76.64	86.60	**87.55**

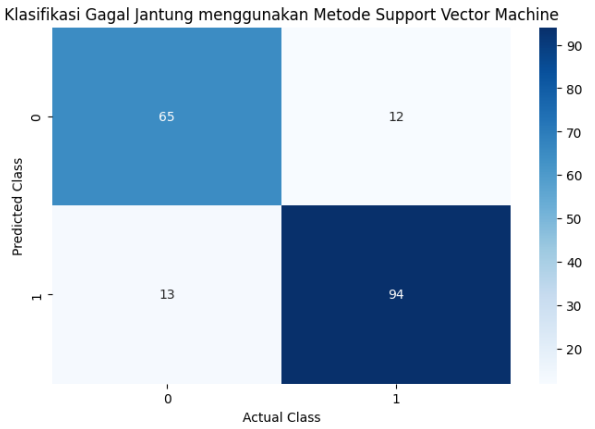
Tabel 6. Hasil Uji Coba Kernel Polinomial

Kernel		Polinomial				
C		0.01	0.1	1	10	100
Hasil Evaluasi	Akurasi (%)	54.63	64.71	77.11	84.20	**85.83**
	Presisi (%)	54.63	66.82	83.57	87.80	**87.59**
	Recall (%)	100	70.32	72.32	82.54	**86.28**
	F1-Score (%)	70.66	68.53	77.54	85.09	**86.93**

Hasil uji coba pada Tabel 3 sampai Tabel 5 menunjukkan bahwa semakin banyak *cost*, maka nilai akurasi semakin tinggi. Akurasi tertinggi pada kernel linear, RBF dan polinomial masing-masing sebesar 86.92%, 86.24% dan 85.83%. Kernel yang menghasilkan akurasi paling tinggi, yaitu kernel linear dengan *cost* 1, 10 dan 100. Hal tersebut dikarenakan kernel linear cenderung lebih sederhana dan memiliki risiko *overfitting* yang lebih rendah, sehingga dapat melakukan generalisasi lebih baik pada data baru yang tidak terlihat. Kernel RBF dan polinomial cenderung menciptakan model yang lebih kompleks. Apabila data tidak memiliki pola *non-linear* yang signifikan, model tersebut dapat menyebabkan *overfitting* pada data pelatihan. Hal ini dapat menangkap *noise* daripada pola yang sebenarnya.

Proses pengujian dengan hasil terbaik menggunakan *cost* 100 menghasilkan presisi, F1-*score*, akurasi dan *recall* masing-masing sebesar 88.68%, 88.26%, 86.41% dan 87.85%. Pengujian

tersebut divisualisasikan menggunakan *Confusion Matrix* yang ditunjukkan pada Gambar 4.



Gambar 3. Hasil *Confusion Matrix*

`` Pada Gambar 4 menunjukkan hasil pengujian dari klasifikasi gagal jantung dengan jumlah TP, TN, FP dan FN berturut-turut sebanyak 65, 94, 12 dan 13. Berdasarkan hasil dari *Confusion Matrix* tersebut, penelitian tentang klasifikasi penyakit gagal jantung menggunakan SVM dapat membantu tenaga medis mendeteksi tanda-tanda awal gagal jantung lebih cepat dan lebih andal. Deteksi dini penyakit ini memungkinkan penanganan lebih awal, sehingga dapat meningkatkan hasil pengobatan dan mengurangi tingkat kematian.

4. KESIMPULAN

Berdasarkan hasil penelitian, diperoleh kesimpulan bahwa metode *Support Vector Machine* cukup baik dalam melakukan klasifikasi pada penyakit gagal jantung. Pada penelitian ini, uji coba yang dilakukan memperoleh nilai akurasi tertinggi sebesar 86.92% menggunakan kernel linear dan *cost* 1, 10 dan 100. Berdasarkan perolehan model terbaik, proses pengujian menghasilkan presisi, F1-*score*, akurasi dan *recall* masing-masing sebesar 88.68%, 88.26%, 86.41% dan 87.85% dengan *cost* 100. Hasil tersebut menunjukkan bahwa atribut yang digunakan dalam model mungkin tidak semuanya relevan untuk klasifikasi penyakit gagal jantung. Beberapa atribut memberikan sedikit atau tidak ada informasi yang berguna untuk prediksi, yang dapat menyebabkan hasil kurang komprehensif. Saran yang dapat disampaikan adalah menambahkan kernel sigmoid yang berperan sebagai fungsi aktivasi dalam jaringan saraf, nilai gamma pada kernel RBF untuk menentukan seberapa jauh pengaruh satu sampel pelatihan, dan nilai degree pada kernel polinomial untuk mengontrol kompleksitas model.

DAFTAR PUSTAKA

[1] F. Syamsuddin, A. Ayuba, and N. I. Nasir, "The Effect of Cardiac Diet Counseling on Knowledge of Heart Diet in Congestive Heart Failure (CHF) Patients at Prof.Dr.H Aloei Saboe Hospital, Gorontalo City," *J. Community Heal. Provis.*, vol. 2, no. 1, pp. 35–41, 2022, doi: 10.55885/jchp.v2i1.116.

[2] K. K. RI, "Penyakit Jantung Penyebab Utama Kematian, Kemenkes Perkuat Layanan Primer," *Public*, 2022.

[3] R. Yunus, U. Ulfa, and M. D. Safitri, "Application of the K-Nearest Neighbors (K-NN) Algorithm for Classification of Heart Failure," *J. Appl. Intell. Syst.*, vol. 6, no. 1, pp. 1–9, 2021.

[4] D. Puspita and M. Fadil, "Penggunaan Ventilasi Mekanik pada Gagal Jantung Akut," *J. Kesehat. Andalas*, vol. 9, no. 1S, pp. 194–203, 2020, doi: 10.25077/jka.v9i1s.1172.

[5] World Health Organization (WHO), "Heart Disease," 2019.

[6] L. Hasibuan, "Ini 4 Penyebab Utama Penyakit Jantung di Indonesia," *CNBC Indonesia*, 2022.

[7] R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, and A. A. Basha, "Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform," *Meas. J. Int. Meas. Confed.*, vol. 146, pp. 800–805, 2019, doi: 10.1016/j.measurement.2019.05.083.

[8] S. Styawati and K. Mustofa, "A Support Vector Machine-Firefly Algorithm for Movie Opinion Data Classification," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 13, no. 3, p. 219, 2019, doi: 10.22146/ijccs.41302.

[9] L. Mohan, J. Pant, P. Suyal, and A. Kumar, "Support Vector Machine Accuracy Improvement with Classification," *Proc. - 2020 12th Int. Conf. Comput. Intell. Commun. Networks, CICN 2020*, pp. 477–481, 2020, doi: 10.1109/CICN49253.2020.9242572.

[10] C. Aroef, Y. Rivan, and Z. Rustam, "Comparing random forest and support vector machines for breast cancer classification," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 18, no. 2, pp. 815–821, 2020, doi: 10.12928/TELKOMNIKA.V18I2.14785.

[11] D. Hsu, V. Muthukumar, and J. Xu, "On the proliferation of support vectors in high dimensions," *Proc. Mach. Learn. Res.*, vol. 130, pp. 91–99, 2021.

[12] M. A. Chandra and S. S. Bedi, "Survey on SVM and their application in image classification," *Int. J. Inf. Technol.*, vol. 13, no. 5, 2021, doi: 10.1007/s41870-017-0080-1.

[13] W. Dudzik, J. Nalepa, and M. Kawulok, "Evolving data-adaptive support vector machines for binary classification," *Knowledge-Based Syst.*, vol. 227, p. 107221, 2021, doi: 10.1016/j.knosys.2021.107221.

[14] D. A. Anggoro and N. D. Kurnia, "Comparison of Accuracy Level of Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) Algorithms in Predicting Heart Disease," vol. 8, no. 5, 2020.

[15] T. J. Osborn *et al.*, "Land Surface Air Temperature Variations Across the Globe Updated to 2019: The CRUTEM5 Data Set," *J. Geophys. Res. Atmos.*, vol. 126, no. 2, 2021, doi: 10.1029/2019JD032352.

[16] D. Mustafa Abdullah and A. Mohsin Abdulazeez, "Machine Learning Applications based on SVM Classification A Review," *Qubahan Acad. J.*, vol. 1, no. 2, pp. 81–90, 2021, doi: 10.48161/qaj.v1n2a50.

[17] M. Tanveer, T. Rajani, R. Rastogi, Y. H. Shao, and M. A. Ganaie, "Comprehensive review on twin support vector machines," *Ann. Oper. Res.*, 2022, doi: 10.1007/s10479-022-04575-w.

[18] S. Styawati, A. Nurkholis, A. A. Aldino, S. Samsugi, E. Suryati, and R. P. Cahyono, "Sentiment Analysis on Online Transportation Reviews Using Word2Vec Text Embedding Model Feature Extraction and Support Vector Machine (SVM) Algorithm," *2021 Int. Semin. Mach. Learn. Optim. Data Sci. ISMODE 2021*, no. January, pp. 163–167, 2022, doi: 10.1109/ISMODE53584.2022.9742906.

[19] K. G. Kapoor and A. Kumar, "SVM Hyperplane Misclassification Control by Finding Optimum Cost of Misclassification with Boundary Value Analysis Technique," vol. 6, no. 3, pp. 2812–2816, 2022.

[20] M. Thohir, A. Z. Foeady, D. C. R. Novitasari, A. Z. Arifin, B. Y. Phiadelvira, and A. H. Asyhar, "Classification of Colposcopy Data Using GLCM-SVM on Cervical Cancer," *2020 Int. Conf. Artif. Intell. Inf. Commun. ICAIIC 2020*, pp. 373–378, 2020, doi: 10.1109/ICAIIC48513.2020.9065027.

[21] Víctor Blanco, A. Japón, and J. Puerto, "A mathematical programming approach to SVM-based classification with label noise," *Comput. Ind. Eng.*, vol. 172, no. 1, 2022, doi: http://dx.doi.org/10.1016/j.cie.2022.108611.

[22] R. Guido, M. C. Groccia, and D. Conforti, "A hyper-parameter tuning approach for cost-sensitive support vector machine classifiers," *Soft Comput.*, vol. 27, no. 18, pp. 12863–12881, 2023, doi: 10.1007/s00500-022-06768-8.

[23] A. Tharwat, "Parameter investigation of support vector machine classifier with kernel functions," *Knowl. Inf. Syst.*, vol. 61, no. 3, pp. 1269–1302, 2019, doi: 10.1007/s10115-019-01335-4.

[24] A. Gilardi, R. Borgoni, and J. Mateu, "A non-separable first-order spatio-temporal intensity for events on linear networks: an application to ambulance interventions," pp. 1–31, 2021.

[25] B. Wang, X. Zhang, S. Xing, C. Sun, and X. Chen, "Sparse representation theory for support vector machine kernel function selection and its application in high-speed bearing fault diagnosis," *ISA Trans.*, vol. 118, no. xxxx, pp. 207–218, 2021, doi: 10.1016/j.isatra.2021.01.060.

[26] A. P. Gopi, R. N. S. Jyothi, V. L. Narayana, and K. S. Sandeep, "Classification of tweets data based on polarity using improved RBF kernel of SVM," *Int. J. Inf. Technol.*, vol. 15, no. 2, pp. 965–980, 2020, doi: 10.1007/s41870-019-00409-4.

[27] Z. Liu *et al.*, "Instance-Variant Loss with Gaussian RBF Kernel for 3D Cross-modal Retriveal," *Proc. Make sure to enter correct Conf. title from your rights confirmation emai*, vol. 1, no. 1, 2023, [Online]. Available: http://arxiv.org/abs/2305.04239

[28] L. Shi, X. Wang, and Y. Shen, "Research on 3D face recognition method based on LBP and SVM," *Optik (Stuttg).*, vol. 220, p. 165157, 2020, doi: 10.1016/j.ijleo.2020.165157.

[29] S. Heilig, M. Münch, and F.-M. Schleif, "Revisiting Memory Efficient Kernel Approximation: An Indefinite Learning Perspective," no. 1995, 2021.

[30] W. Tuerxun, X. Chang, G. Hongyu, J. Zhijie, and Z. Huajian, "Fault Diagnosis of Wind Turbines Based on a Support Vector Machine Optimized by the Sparrow Search Algorithm," *IEEE Access*, vol. 9, pp. 69307–69315, 2021, doi: 10.1109/ACCESS.2021.3075547.

[31] V. M. Nguyen-Thanh, C. Anitescu, N. Alajlan, T. Rabczuk, and X. Zhuang, "Parametric deep energy approach for elasticity accounting for strain gradient effects," *Comput. Methods Appl. Mech. Eng.*, vol. 386, no. December, 2021, doi: 10.1016/j.cma.2021.114096.

[32] W. Sadewo, Z. Rustam, H. Hamidah, and A. R. Chusmarsyah, "Pancreatic cancer early detection using twin support vector machine based on kernel," *Symmetry (Basel).*, vol. 12, no. 4, pp. 6–13, 2020, doi: 10.3390/SYM12040667.

[33] E. Westphal and H. Seitz, "A Machine Learning Method for Defect Detection and

Visualization in Selective Laser Sintering Based on Convolutional Neural Networks," *Addit. Manuf.*, vol. 41, no. March, 2021, doi: 10.1016/j.addma.2021.101965.

[34] D. Chicco and G. Jurman, "The Advantages of the Matthews Correlation Coefficient (MCC) Over F1 Score and Accuracy in Binary Classification Evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1–13, 2020, doi: 10.1186/s12864-019-6413-7.

[35] H. Yang *et al.*, "Risk Prediction of Diabetes: Big Data Mining with Fusion of Multifarious Physical Examination Indicators," *Inf. Fusion*, vol. 75, no. March, pp. 140–149, 2021, doi: 10.1016/j.inffus.2021.02.015.

[36] S. Mishra, P. K. Mallick, L. Jena, and G. S. Chae, "Optimization of Skewed Data using Sampling-Based Preprocessing Approach," *Front. Public Heal.*, vol. 8, no. July, pp. 1–7, 2020, doi: 10.3389/fpubh.2020.00274.

[37] M. Shorfuzzaman and M. S. Hossain, "MetaCOVID: A Siamese neural network framework with contrastive loss for n -shot diagnosis of COVID-19 patients," *ELSEVIER*, no. January, 2020.

[38] L. Qadrini, "Undersampling dan K-Fold Random Forest Untuk Klasifikasi Kelas Tidak Seimbang," *Build. Informatics, Technol. Sci.*, vol. 4, no. 4, pp. 1967–1974, 2023, doi: 10.47065/bits.v4i4.3141.

[39] L. Sari, A. Romadloni, and R. Listyaningrum, "Penerapan Data Mining dalam Analisis Prediksi Kanker Paru Menggunakan Algoritma Random," *Infotekmesin*, vol. 14, no. 01, pp. 155–162, 2023, doi: 10.35970/infotekmesin.v14i1.1751.

[40] C. A. C. F. Tram, "Label-Only Membership Inference Attacks," 2021.

[41] N. Fiorentini and M. Losa, "Handling Imbalanced Data in Road Crash Severity Prediction by Machine Learning Algorithms," *Infrastructures*, vol. 5, no. 7, 2020, doi: 10.3390/infrastructures5070061.