



Analisis Deteksi Dini Penyakit Jantung dengan Pendekatan Support Vector Machine pada Data Pasien

INFO PENULIS	INFO ARTIKEL
Fitra M.Natsir Universitas Muhammadiyah Makassar 105841105420@student.unismuh.ac.id	ISSN: 3026-3603 Vol. 2, No. 2 Oktober 2024 http://jurnal.ardenjaya.com/index.php/ajst
Riski Yusliana Bakti Universitas Muhammadiyah Makassar	
Titin Wahyuni Universitas Muhammadiyah Makassar	

© 2024 Arden Jaya Publisher All rights reserved

Saran Penulisan Referensi:

Natsir, F. M., Bakti, R. Y., & Wahyuni, T., (2024). Analisis Deteksi Dini Penyakit Jantung Dengan Pendekatan Support Vector Machine Pada Data Pasien. *Arus Jurnal Sains dan Teknologi*, 2 (1), 437-446.

Abstrak

Penelitian ini mengevaluasi deteksi dini risiko penyakit jantung menggunakan algoritma *Support Vector Machine* (SVM) pada data pasien RSUD Haji Makassar. Dari 640 data pasien yang diproses, tersisa 640 data yang siap untuk analisis. Metode Elbow digunakan untuk mengidentifikasi tiga kluster utama berdasarkan tingkat risiko penyakit jantung: rendah, tinggi, dan sangat tinggi. Model SVM diterapkan dengan akurasi sebesar 0.984375, menunjukkan performa unggul dalam mengklasifikasikan risiko penyakit jantung dengan kesalahan minimal. Analisis ini menegaskan bahwa SVM adalah alat yang efektif untuk identifikasi awal faktor risiko penyakit jantung, dengan pemilihan dan pengolahan data yang tepat berperan krusial dalam meningkatkan akurasi. Temuan ini memberikan dasar bagi pengembangan sistem deteksi dini yang lebih akurat dan intervensi pencegahan yang lebih efektif.

Kata Kunci: Support Vector Machine, klasifikasi data, RSUD Haji Makassar.

Abstract

This study evaluates early detection of heart disease risk using the Support Vector Machine (SVM) algorithm on patient data from RSUD Haji Makassar. Out of 640 patient records processed, 640 data points remained for analysis. The Elbow method was used to identify three main clusters based on heart disease risk levels: low, high, and very high. The SVM model was applied with an accuracy of 0.984375, demonstrating excellent performance in classifying heart disease risk with minimal errors. This analysis underscores that SVM is an effective tool for early identification of heart disease risk factors, with proper data selection and processing playing a crucial role in improving accuracy. These findings provide a foundation for developing more accurate early detection systems and more effective preventive interventions.

Keywords: Support Vector Machine, data klasifikasi, RSUD Haji Makassar.

A. Pendahuluan

Penyakit Jantung merupakan sebutan umum yang digunakan untuk menggambarkan gangguan terhadap fungsi kerja jantung. Penyakit atau gangguan jantung sendiri memiliki banyak jenis dan macam nama penyakitnya seperti kardiovaskuler, jantung koroner dan serangan jantung. Penyakit Jantung merupakan salah satu penyakit yang paling sering terjadi kasusnya pada kalangan masyarakat, dimana penyakit jantung ini dapat menimpa dan menyerang siapapun tanpa memandang usia, jenis kelamin dan gaya hidup. Menurut WHO (Organisasi Kesehatan Dunia) dan CDC, penyakit jantung adalah penyebab utama kematian di Inggris, Amerika Serikat, Kanada dan Australia. Jumlah orang dewasa yang didiagnosis dengan penyakit jantung terdiri dari 26,6 Juta Jiwa (11,3%) dari populasi orang dewasa (Utomo & Mesran, 2020).

Data telah menunjukkan bahwa penyakit jantung merupakan penyakit yang mematikan, sehingga perlu kewaspadaan dari setiap orang akan resiko penyakit jantung ini. Salah satu kontribusi dalam bidang teknologi untuk memberikan informasi bermanfaat mengenai resiko penyakit jantung adalah dengan pengolahan data atau teknik data mining. Dalam data mining ada beberapa pendekatan seperti klusterisasi, klasifikasi, asosiasi. Pada penelitian ini dilakukan teknik klasifikasi yaitu pengelompokan data, data yang dikelompokkan berlandaskan relasi atau hubungan data terhadap data sampel (Pradana et al., 2022).

Jantung adalah salah satu rongga organ berotot yang terdapat di dalam tubuh seluruh manusia dan berperan sangat penting karena sangat berperan dalam sistem peredaran darah dan juga sebagai pemompa dan penerima darah ke seluruh tubuh (Lailatul Mufidah, 2021).

Penyakit jantung adalah suatu kondisi di mana bagian jantung, seperti pembuluh darah jantung, lapisan jantung, katup jantung, dan otot jantung menjadi tidak berfungsi, penyakit ini dapat memiliki banyak penyebab, seperti penyumbatan di arteri jantung, peradangan, infeksi, atau cacat lahir (Hidayat et al., 2023). Abrasi merupakan suatu proses alam yang terjadi berupa pengikisan daratan pantai yang disebabkan oleh gelombang dan arus laut yang merusak atau biasa juga disebut erosi pantai (Setiyono 1996).

Pengenalan dini penyakit jantung bawaan menjadi penting karena presentasi klinis dan kemunduran kondisi umum yang mendadak dapat menimbulkan resiko morbiditas, dan mortalitas yang sebenarnya dapat dihindari (Eva Miranda Marwali et al., 2021).

Deteksi dini dan penanganan awal hipertensi dan faktor resiko lainnya, telah berkontribusi secara bertahap terhadap turunya kematian karena serangan jantung dan stroke di negara berpenghasilan tinggi lebih dari tiga decade terakhir (Widiastuti et al., 2021)

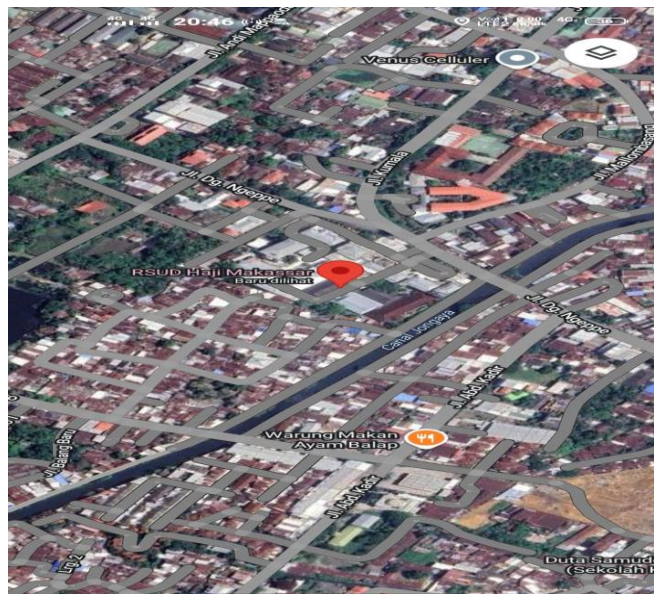
Deteksi dini dan penanganan empat faktor resiko mayor, yaitu merokok, hipertensi, diabetes melitus, dan hiperkolesterolemia sangat penting sebagai upaya pencegahan primer dan sekunder terhadap penyakit kardiovaskuler. Penelitian yang dilakukan oleh Cohn, et al (2003)

Pasien adalah seseorang yang menerima perawatan medis. Sering kali, pasien menderita penyakit atau cedera dan memerlukan bantuan dokter untuk memulihkannya. Kata pasien dari bahasa Indonesia analog dengan kata patient dari bahasa Inggris (Rochman et al., 2019).

Setelah melakukan *studi* literatur pada penelitian sebelumnya, yang berjudul “PENGUATAN PERAN KADER KESEHATAN DALAM DETEKSI DINI RESIKO PENYAKIT JANTUNG DI KELURAHAN TAMBAKREJA KABUPATEN CILACAP” (Kasron et al., 2022), maka tentu penelitian ini memiliki perbedaan dari penelitian terdahulu sebagai penelitian yang mendukung penelitian ini. Berdasarkan latar belakang tersebut, tujuan penelitian ini untuk melakukan klasifikasi data dengan baik dengan judul “ANALISIS DETEKSI DINI PENYAKIT JANTUNG DENGAN PENDEKATAN *SUPPORT VECTOR MACHINE* PADA DATA PASIEN”.

B. Metodologi

Lokasi pengambilan data pasien pada penelitian ini dilakukan di Rumah Sakit Umum Daerah (RSUD) Haji Makassar di Pintu keluar rs haji, Jl. Dg. Ngeppe, Balang Baru, Kec. Tamalate, Kota Makassar, Sulawesi Selatan 90122. Penelitian ini dijadwalkan akan dilaksanakan dari bulan Mei 2024 hingga Agustus 2024.



Gambar 2. Peta lokasi penelitian

a. Teknik Pengujian Sistem

Teknik pengujian sistem melibatkan pembagian data menjadi set pelatihan dan set pengujian (train-test split). Model SVM dilatih menggunakan set pelatihan, kemudian diuji pada set pengujian. Kinerja model dievaluasi melalui berbagai metrik, seperti akurasi, sensitivitas, dan spesifisitas, untuk menentukan seberapa baik model mengidentifikasi faktor risiko penyakit kronis. Validasi kinerja ini memastikan bahwa model tidak hanya efektif pada data pelatihan tetapi juga pada data yang belum pernah dilihat sebelumnya.

b. Teknis Analisis Data

Analisis data meliputi pembersihan, normalisasi, dan seleksi fitur melalui *Principal Component Analysis* (PCA) atau metode berbasis model. Data dibagi untuk pelatihan dan pengujian, lalu model SVM dilatih. Evaluasi kinerja dilakukan dengan metrik seperti akurasi dan F1 Score, serta cross-validation jika diperlukan. Hasil klasifikasi dianalisis untuk memahami risiko penyakit kronis dan kontribusi fitur.

C. Hasil dan Pembahasan

a. Analisa Data Mentah

Data yang diolah dalam penelitian ini berdasarkan data pasien di RSUD Haji Makassar, periode 1 Januari 2021 sampai 15 Juli 2024. Dataset ini terdiri dari berbagai atribut seperti NoRM, NoPen, Nama Pasien, JK, Tanggal Lahir, Umur, Tanggal Registrasi, Unit Pelayanan, Dokter, Tanggal Keluar, dan Ruang Akhir. Dataset terdiri dari 640 data pasien yang kemudian data tersebut akan dilakukan *preprocessing* data. sebagai berikut :

No.	No. RM	NO PEN	Nama Pasien	JK	Tgl Lahir/Umur	Tgl Reg	Unit Pelayanan	Dokter	Tgl Keluar	Ruang Akhir
1	64424	2109200005	***	P	08-10-1977 (43 Th/ 11 bl/ 12 hr)	20-09-2021 08:37:25	Poli Jantung dan Pembuluh Darah	DR AMELIA ARINDANIE SYAHRIR SPJP FIHA	22-09-2021 16:28:43	Poli Jantung dan Pembuluh Darah
2	67738	2109150031	***	P	05-12-1945 (75 Th/ 9 bl/ 10 hr)	15-09-2021 09:30:23	Poli Jantung dan Pembuluh Darah	DR AMELIA ARINDANIE SYAHRIR SPJP FIHA	16-09-2021 08:28:00	Poli Jantung dan Pembuluh Darah
3	91996	2109150014	***	L	09-02-1964 (57 Th/ 7 bl/ 6 hr)	15-09-2021 08:41:44	Poli Jantung dan Pembuluh Darah	DR AMELIA ARINDANIE SYAHRIR SPJP FIHA	16-09-2021 08:18:47	Poli Jantung dan Pembuluh Darah
4	100402	2109290012	***	P	05-06-1979 (42 Th/ 3 bl/ 24 hr)	29-09-2021 08:30:11	Poli Jantung dan Pembuluh Darah	DR AMELIA ARINDANIE SYAHRIR SPJP FIHA	30-09-2021 11:18:24	Poli Jantung dan Pembuluh Darah
5	123244	2109200045	***	L	29-12-1964 (56 Th/ 8 bl/ 22 hr)	20-09-2021 10:02:03	Poli Jantung dan Pembuluh Darah	DR AMELIA ARINDANIE SYAHRIR SPJP FIHA	23-09-2021 09:59:54	Poli Jantung dan Pembuluh Darah
...
640	227838	2407150089	***	L	07-10-1950 (73 Th/ 9 bl/ 8 hr)	15-07-2024 10:16:52	Poli Jantung dan Pembuluh Darah	DR AMELIA ARINDANIE SYAHRIR SPJP FIHA		Poli Jantung dan Pembuluh Darah

b. Data Preprocessing

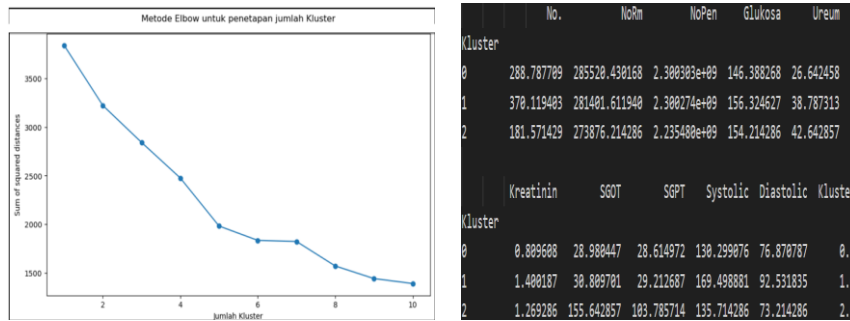
Proses pra-pemrosesan data bertujuan memastikan data bersih dan siap untuk analisis SVM, dimulai dengan pengumpulan data pasien dari RSUD Haji Makassar. Setelah menghapus data yang hilang atau tidak valid dan memilih fitur relevan, diperoleh 640 data pasien yang digunakan sebagai data uji. Berikutnya, tekanan darah dibagi menjadi systolic dan diastolic, kemudian data dikelompokkan berdasarkan kemiripan untuk mengklaster data ke dalam grup dengan karakteristik yang sama.

No	No Rm	NoPen	Nama Pasien	JK	Tanggal lahir	Usia	Tanggal Registrasi	Unit Pelayanan	Dokter	Tanggal Keluar	Ruang Akhir	Glu kosa	Ur eum	Krea tinin	SG OT	SG PT	Tekana nDarah	Pengli hatan	Pende ngaran	Penci uman	Bic ara	Perna fasan
1	285 519	22071 90085	***	L	12/31/ 1948	73	7/19/202 2 11:09	Poli Jantun g dan Pembul uh Darah	dr. ANDI MUHA MMAD REIS R SAIBY, Sp.JP	19-07- 2022 12:12: 32	Poli Jantu ng dan Pemb uluh Darah	135	21	0.57	23	12	119/65	Normal	Normal	Normal	Normal	Gang guan
2	289 769	22030 20059	***	P	5/12/1 970	51	2/3/2022 10:38	Poli Jantun g dan Pembul uh Darah	DR AMELIA ARIND ANIE SYAH RIR SPJP FIHA	9/3/20 22 11:36	Poli Jantu ng dan Pemb uluh Darah	93	18	0.76	17	15	130/90	Normal	Normal	Normal	Normal	Normal
640	295 262	24050 30022	***	P	12/31/ 1962	61	3/5/2024 8:10	Poli Jantun g dan Pembul uh Darah	dr. ANDI MUHA MMAD REIS R SAIBY, Sp.JP	3/5/20 24 10:41	Poli Jantu ng dan Pemb uluh Darah	168	50	0.53	42	18	129/78	Normal	Normal	Normal	Normal	Normal

Gambar 4. Dataset Pasien RSUD Haji Makassar pra pemrosesan

c. Implementasi Metode Elbow dan Menentukan Hasil Clustering

Metode Elbow digunakan untuk menentukan jumlah cluster optimal dengan memplot Sum of Squared Errors (SSE) terhadap jumlah cluster, yang menunjukkan titik elbow pada tiga cluster. Klasterisasi K-Means membagi data pasien menjadi tiga kluster berdasarkan tekanan darah: rendah, tinggi, dan sangat tinggi, memfasilitasi identifikasi pola risiko.



Gambar 5. Hasil Metode Elbow dan Output Hasil Clustering

d. Implementasi Model SVM

Model SVM menunjukkan kinerja sangat baik dengan akurasi 0.984375 atau 98.44%, mengklasifikasikan sebagian besar data dengan benar. Laporan klasifikasi menunjukkan *precision*, *recall*, dan *f1-score* yang hampir sempurna untuk setiap kelas, dengan performa terbaik pada kelas 2. Rata-rata macro avg untuk *precision*, *recall*, dan *f1-score* adalah 0.99, menandakan performa konsisten di seluruh kelas, sementara weighted avg adalah 0.98, mencerminkan kinerja model yang solid. Secara keseluruhan, model SVM efektif dalam klasifikasi dengan tingkat kesalahan minimal.

Tabel 1. Hasil Akurasi

Akurasi: 0.984375		Laporan Klasifikasi:		
	Precision	recall	f1-score	support
0	0.97	1.00	0.99	77
1	1.00	0.96	0.98	49
2	1.00	1.00	1.00	2
Accuracy			0.98	128
Macro avg	0.99	0.99	0.99	128
Weighted avg	0.98	0.98	0.98	128

Penjelasan mengenai metrik-metrik dalam laporan klasifikasi:

- Model SVM: SVM bekerja dengan mencari *hyperplane* optimal yang memisahkan kelas dalam ruang fitur dengan margin maksimum, meningkatkan kemampuan model untuk menggeneralisasi data yang belum terlihat. Akurasi model SVM yang diperoleh adalah 0.984375, artinya dari 128 sampel data pengujian, model berhasil memprediksi 126 sampel dengan baik.
- Laporan klasifikasi: Model SVM yang digunakan berhasil mengklasifikasikan data dengan sangat baik, seperti yang terlihat dari metrik *precision*, *recall*, dan *f1-score* di setiap kelas.

Akurasi:

Akurasi 0.984375 (98.44%): artinya model ini berhasil mengklasifikasikan 126 dari 128 sampel data dengan benar. Akurasi yang tinggi ini menunjukkan kinerja model yang sangat baik, tetapi penting juga memperhatikan metrik lain seperti *precision*, *recall*, dan *f1-score*, terutama jika ada ketidakseimbangan kelas. Metrik tambahan ini memastikan bahwa akurasi tinggi benar-benar mencerminkan kemampuan model dalam mengklasifikasikan semua kelas secara konsisten.

Evaluasi Per Kelas:

Kelas 0

- a. **Precision (0.97):** Dari semua prediksi model yang menyatakan data termasuk kelas 0, 97% di antaranya benar-benar merupakan kelas 0, menunjukkan ketepatan model dalam mengklasifikasikan kelas ini.
- b. **Recall (1.00):** Model berhasil mengidentifikasi semua contoh kelas 0 dengan benar, tanpa ada yang terlewat, menunjukkan sensitivitas model yang sempurna untuk kelas ini.
- c. **F1-Score (0.99):** Kombinasi harmonis antara *precision* dan *recall* menghasilkan f1-score 0.99, menandakan keseimbangan yang sangat baik antara ketepatan dan sensitivitas model dalam mengklasifikasikan kelas 0.
- d. **Support (77):** Ada 77 sampel data yang termasuk dalam kelas 0, yang digunakan untuk menghitung metrik-metrik tersebut.

Kelas 1

- a) **Precision (1.00):** Model berhasil mengklasifikasikan semua data yang diprediksi sebagai kelas 1 dengan benar, menunjukkan bahwa tidak ada kesalahan dalam prediksi untuk kelas ini.
- b) **Recall (0.96):** Dari semua data yang sebenarnya termasuk dalam kelas 1, model berhasil mengidentifikasi 96% di antaranya, dengan sedikit data yang terlewat.
- c) **F1-Score (0.98):** Menunjukkan keseimbangan yang sangat baik antara *precision* dan *recall*, artinya model memiliki kinerja yang hampir sempurna dalam mengklasifikasikan kelas 1.
- d) **Support (49):** Ada 49 sampel data yang termasuk dalam kelas 1, yang menjadi dasar perhitungan metrik-metrik tersebut.

Kelas 2

- a. **Precision (1.00):** Model berhasil mengklasifikasikan semua data yang diprediksi sebagai kelas 2 dengan benar, tanpa ada kesalahan dalam prediksi.
- b. **Recall (1.00):** Semua data yang sebenarnya termasuk dalam kelas 2 berhasil diidentifikasi oleh model, menunjukkan sensitivitas yang sempurna.
- c. **F1-Score (1.00):** Menunjukkan keseimbangan sempurna antara *precision* dan *recall*, menandakan kinerja yang ideal dalam mengklasifikasikan kelas 2.
- d. **Support (2):** Hanya ada 2 sampel data yang termasuk dalam kelas 2, yang digunakan untuk menghitung metrik-metrik tersebut.

Rata-Rata

Macro avg (Rata-Rata makro): *Macro average* adalah rata-rata sederhana dari *precision*, *recall*, dan f1-score di semua kelas, tanpa memperhitungkan jumlah sampel di setiap kelas. Dalam kasus ini, *macro avg* untuk *precision*, *recall*, dan f1-score adalah 0.99, yang menunjukkan performa model yang konsisten dan sangat baik di semua kelas, meskipun ada perbedaan jumlah sampel di masing-masing kelas.

- a) **Macro Avg Precision: 0.99**
- b) **Macro Avg Recall: 0.99**
- c) **Macro Avg F1-Score: 0.99**

Weight avg (Rata-Rata berbobot): *Weighted average* memperhitungkan jumlah sampel di setiap kelas saat menghitung rata-rata *precision*, *recall*, dan f1-score. Dalam hasil ini, *weighted avg* untuk *precision*, *recall*, dan f1-score adalah 0.98. Ini menunjukkan bahwa, meskipun beberapa kelas memiliki lebih sedikit data, model tetap memiliki performa yang kuat secara keseluruhan, dengan metrik yang mencerminkan kontribusi proporsional dari setiap kelas sesuai dengan ukuran data mereka.

- a. **Weighted Avg Precision: 0.98**
- b. **Weighted Avg Recall: 0.98**
- c. **Weighted Avg F1-Score: 0.98**

Tabel 2. Hasil Prediksi

N	Nam	Gl	Ur	Kre	S	S	Pen	Pend	Pen	Bi	Per	Sy	Dia	Kl	Predi
o	a	uk	eu	ati	G	G	glih	enga	ciu	ca	nafa	st	sto	as	ksiKl
	Pasi	os	m	nin	O	P	atan	ran	ma	ra	san	oli	lic	te	aster
	en	a			T	T			n			c		r	

3	***	26 9	21	0.7 1	9 5	8 3	Nor mal	Nor mal	Nor mal	No rm al	Nor mal	15 0	45	2	Sanga t Tinggi
7	***	12 5	27	0.8 6	8 4	4 7	Nor mal	Nor mal	Nor mal	No rm al	Nor mal	15 1	58	0	Renda h
1 1	***	15 8	23	1.2 7	2 2	2 7	Nor mal	Nor mal	Nor mal	No rm al	Gan gua n	14 0	90	0	Renda h
...
6 3 9	***	15 2	69	1.1 2	1 3	1 5	Nor mal	Nor mal	Nor mal	No rm al	Nor mal	16 0.7		1	Tinggi

Model SVM memberikan kinerja yang sangat baik dengan akurasi tinggi sebesar 98.44%. Laporan klasifikasi menunjukkan *precision*, *recall*, dan *f1-score* yang hampir sempurna untuk setiap kelas, dengan kelas 2 diklasifikasikan secara akurat 100%. Rata-rata *macro avg* untuk *precision*, *recall*, dan *f1-score* adalah 0.99, menandakan performa yang konsisten dan sangat baik di seluruh kelas, sementara *weighted avg* menunjukkan kinerja model yang solid dengan nilai 0.98. Secara keseluruhan, model SVM menunjukkan kemampuan yang kuat dalam mengidentifikasi dan mengklasifikasikan data dengan tingkat kesalahan minimal.

Pengujian Sistem

Pengujian sistem clustering dengan metode Elbow membagi data menjadi tiga klaster risiko penyakit jantung: Klaster 0 (risiko rendah, 358 sampel), Klaster 1 (risiko tinggi, 268 sampel), dan Klaster 2 (risiko sangat tinggi, 14 sampel). Pengujian sistem klasifikasi dan prediksi dengan model SVM melibatkan pelatihan untuk menemukan hyperplane optimal yang memisahkan kelas data, diikuti dengan evaluasi menggunakan data uji. Evaluasi mencakup akurasi, presisi, recall, *f1-score*, serta *macro* dan *weighted average* untuk menilai kinerja model secara menyeluruh. Hasil pengujian menunjukkan efektivitas model dalam mengidentifikasi dan mengklasifikasikan risiko penyakit jantung. Model SVM memberikan gambaran mendetail tentang kinerja klasifikasi dan akurasi.

Pengujian sistem *clustering elbow* dengan melihat hasil klaster pada kolom klaster dimana nilai 0 itu mempresentasikan resiko rendah penyakit jantung tinggikan nilai 1 mempresentasikan resiko tinggi penyakit jantung dan nilai 2 mempresentasikan resiko sangat tinggi penyakit jantung.

Tabel 3. Clustering

N o	Nam a Pasi en	U si a	Gl uk osa	Ur eu m	Kre ati nin	S G O T	S G P T	Teka nanD arah	Pen glin atan	Pend enga ran	Pen ciu man	Bi ca ra	Per nafa san	Sy sto lic	Dia sto lic	Kl as te r
1	***	7 3	13 5	21	0.5 7	2 3	1 2	119/ 65	Nor mal	Nor mal	Nor mal	No rm al	Gan gua n	11 9	65	0
2	***	5 1	93	18	0.7 6	1 7	1 5	130/ 90	Nor mal	Nor mal	Nor mal	No rm al	Nor mal	13 0	90	0
...
6 4 0	***	6 1	16 8	50	0.5 3	4 2	1 8	129/ 78	Nor mal	Nor mal	Nor mal	No rm al	Nor mal	12 9	78	0

Selanjutnya mencari jumlah data klaster beresiko tinggi dan beresiko rendah dari total 640 data pasien deteksi dini.

Tabel 4. Jumlah Prediksi

No	Nama Pasien	Usia	Glukosa	Ureum	Kreatinin	SGOT	SGPT	Tekanan Darah	Penyglihan	Pendengaran	Penycairan	Bicafasan	Pernafasan	Syostolic	Diaostolic	Klaster
1	***	73	135	21	0.57	23	12	119/65	Nor	Nor	Nor	Normal	Gangguan	119	65	0
2	***	51	93	18	0.76	17	15	130/90	Nor	Nor	Nor	Normal	Nor	130	90	0
...
640	***	61	168	50	0.53	42	18	129/78	Nor	Nor	Nor	Normal	Nor	129	78	0
JUMLAH PREDIKSI KLASTER 0															358	
JUMLAH PREDIKSI KLASTER 1															268	
JUMLAH PREDIKSI KLASTER 2															14	

Berdasarkan data hasil prediksi, sistem ini mengidentifikasi tiga klaster berdasarkan risiko penyakit jantung. Klaster 0, dengan 358 sampel, menunjukkan kesehatan relatif normal. Klaster 1, yang berjumlah 268 sampel, mengindikasikan adanya risiko penyakit jantung, sementara klaster 2, dengan 14 sampel, menunjukkan risiko yang lebih tinggi terhadap penyakit jantung. Hal ini telah sesuai dengan yang dijelaskan oleh narasumber dari penelitian ini yaitu Dr. Husnul Khatimah S,Ked. Selaku Dokter Rumah Sakit Umum Daerah (RSUD) Haji Makassar mengatakan bahwa dapat dilihat dari hasil pemeriksaan laboratorium yang menunjukkan bahwa pasien yang mengalami peningkatan kadar ureum dan kreatinin mengidiskasikan adanya kerusakan fungsi ginjal, serta hipertensi yang dapat menyebabkan diabetes juga dapat meningkatkan aterosklerosis. Selain itu kadar SGOT dan SGPT yang tinggi mengindikasikan gangguan pada hati dan jantung pasien.

e. Menghitung Metode Evaluasi

- 1) Confusion Matrix: Confusion matrix adalah tabel yang digunakan untuk mendeskripsikan kinerja model klasifikasi. Tabel ini menunjukkan jumlah prediksi yang benar dan salah untuk setiap kelas:

	Prediksi 0	Prediksi 1	Prediksi 2
KELAS 0	75	2	0
KELAS 1	2	47	0
KELAS 2	0	0	2

- 2) **Akurasi:** Akurasi dihitung sebagai rasio jumlah prediksi yang benar terhadap total prediksi.

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Prediksi}} = \frac{126}{128} = 0.984375$$

- 3) **Precision (Presisi):** Precision mengukur seberapa banyak prediksi positif yang benar dari semua prediksi positif.

$$\text{Presisi}_i = \frac{TP_i}{TP_i + FP_i}$$

a) Kelas 0: $\frac{75}{75+2} = 0.97$

b) Kelas 1: $\frac{49}{49+0} = 1.00$

c) Kelas 2: $\frac{2}{2+0} = 1.00$

d) *Macro Average* adalah rata-rata *Precision* dari semua kelas.

$$\text{Macro Average} = \frac{0.97 + 1.00 + 1.00}{3} = 0.99$$

e) *Weighted Average* menghitung *Precision* dengan mempertimbangkan dukungan (jumlah instance) dari setiap kelas.

$$\text{Weighted Average} = \frac{(77 \times 0.97) + (49 \times 1.00) + (2 \times 1.00)}{128} = 0.98$$

4) **Recall:** Recall mengukur seberapa banyak instance yang benar-benar positif terdeteksi dari semua instance positif aktual. Untuk setiap kelas.

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i}$$

a) Kelas 0: $\frac{75}{75+0} = 1.00$

b) Kelas 1: $\frac{49}{49+2} = 0.96$

c) Kelas 2: $\frac{2}{2+0} = 1.00$

d) *Macro Average* adalah rata-rata *Recall* dari semua kelas.

$$\text{Macro Average} = \frac{1.00 + 0.96 + 1.00}{3} = 0.99$$

e) *Weighted Average* menghitung *Recall* dengan mempertimbangkan dukungan (jumlah instance) dari setiap kelas.

$$\text{Weighted Average} = \frac{(77 \times 1.00) + (49 \times 0.96) + (2 \times 1.00)}{128} = 0.98$$

5) **F1-Score:** adalah rata-rata harmonis dari *precision* dan *recall*, memberikan ukuran keseimbangan antara keduanya.

$$F1 - \text{Score}_i = 2 \times \frac{\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}$$

a) Kelas 0: $2 \times \frac{0.97 \times 1.00}{0.97 + 1.00} = 0.99$

b) Kelas 1: $2 \times \frac{1.00 \times 0.96}{1.00 + 0.96} = 0.98$

c) Kelas 2: $2 \times \frac{1.00 \times 1.00}{1.00 + 1.00} = 1.00$

d) *Macro Average* adalah rata-rata *F1-Score* dari semua kelas.

$$\text{Macro Average} = \frac{0.99 + 0.98 + 1.00}{3} = 0.99$$

e) *Weighted Average* menghitung *F1-Score* dengan mempertimbangkan dukungan (jumlah instance) dari setiap kelas.

$$\text{Weighted Average} = \frac{(77 \times 0.99) + (49 \times 0.98) + (2 \times 1.00)}{128} = 0.98$$

Penelitian ini menganalisis deteksi dini penyakit jantung menggunakan algoritma Support Vector Machine (SVM) pada data pasien dari RSUD Haji Makassar. Dari total 3522 data pasien, dilakukan pembersihan dan pemrosesan, sehingga menghasilkan 640 data yang siap digunakan untuk pengujian. Selanjutnya, metode Elbow diterapkan untuk mengidentifikasi tiga kluster utama yang mewakili tingkat risiko penyakit jantung, yaitu risiko rendah, tinggi, dan sangat tinggi. Proses ini memetakan pasien berdasarkan karakteristik medis seperti tekanan darah dan kadar glukosa, memungkinkan pemisahan pasien ke dalam kluster yang sesuai. Dengan demikian, diperoleh pandangan yang lebih jelas mengenai risiko kesehatan setiap pasien.

Implementasi model SVM dalam penelitian ini menunjukkan hasil yang sangat memuaskan, dengan akurasi sebesar 0.984375 atau 98,44%. Kinerja model diukur menggunakan metrik seperti precision, recall, dan f1-score, yang semuanya menunjukkan performa yang kuat dan konsisten di seluruh kelas. Dalam proses klasifikasi, model ini mampu secara efektif memisahkan pasien ke dalam kelompok risiko yang tepat dengan tingkat kesalahan minimal.

Hasil ini mengindikasikan bahwa SVM merupakan alat yang efektif dalam mengklasifikasikan risiko penyakit jantung berdasarkan data pasien. Melalui analisis ini, diperoleh rekomendasi yang lebih akurat untuk deteksi dini dan manajemen risiko penyakit jantung.

D. Kesimpulan

Berdasarkan perhitungan analisis tinggi gelombang dan abrasi maka, dapat diambil kesimpulan bahwa :

1. Algoritma *Support Vector Machine* (SVM) secara efektif mampu mendeteksi dini risiko penyakit jantung pada pasien. Berdasarkan analisis terhadap data pasien RSUD Haji Makassar, metode Elbow berhasil mengidentifikasi tiga kluster utama yang merepresentasikan tingkat risiko penyakit jantung yang berbeda, yaitu risiko rendah, tinggi, dan sangat tinggi. Dengan akurasi mencapai 0.984375 atau 98,44%, model SVM menunjukkan kinerja unggul dalam mengklasifikasikan data pasien dengan tingkat kesalahan yang sangat rendah. Hal ini mengindikasikan bahwa SVM merupakan alat yang andal dalam identifikasi awal faktor risiko penyakit jantung, yang berperan penting dalam pencegahan dan intervensi dini.
2. Pentingnya proses preprocessing data dan pemilihan fitur yang tepat dalam meningkatkan akurasi model pada analisis data medis juga diperhatikan. Atribut-atribut seperti tekanan darah, kadar glukosa, dan fungsi ginjal terbukti krusial dalam penentuan risiko penyakit jantung. Implementasi SVM pada data. pasien ini tidak hanya memberikan hasil signifikan, tetapi juga dapat menjadi dasar bagi pengembangan sistem

Saran

1. Perluasan Dataset: Kumpulkan data tambahan dari sumber atau periode waktu lain untuk meningkatkan keberagaman dan kualitas model.
2. Eksplorasi Algoritma Alternatif: Evaluasi algoritma klasifikasi lain seperti Random Forest atau Deep Learning untuk membandingkan efektivitas.

E. Referensi

- Aji, P. W. (2023). *Klasifikasi Tingkat Kepositifan Pengidap Penyakit Jantung Dengan Pendekatan Algoritma SVM Berbasis Both Kernel (Linear Dan Polynomial)* (Doctoral dissertation, Universitas Pelita Bangsa).
- Rachmawati, C., Martini, S., & Artanti, K. D. (2021). Analisis Faktor Risiko Modifikasi Penyakit Jantung Koroner Di Rsu Haji Surabaya Tahun 2019. *Media Gizi Kesmas*, 10(1), 47-55.
- Kasron, K., & Susilawati, S. (2022). PENANGANAN PENYAKIT JANTUNG DI KELURAHAN TAMBAKREJA KABUPATEN CILACAP. *JABI: Jurnal Abdimas Bhakti Indonesia*, 3(2), 21-28.
- Lestari, W., & Sumarlinda, S. (2023). Studi Komparatif Model Klasifikasi Kerentanan Penyakit Jantung Menggunakan Algoritma Machine Learning. *SATIN-Sains dan Teknologi Informasi*, 9(1), 107-115.
- Prabowo, A. S., & Kurniadi, F. I. (2023). Analisis Perbandingan Kinerja Algoritma Klasifikasi dalam Mendeteksi Penyakit Jantung. *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, 7(1), 56-61.
- Pradana, M. G., Saputro, P. H., & Wijaya, D. P. (2022). Komparasi Metode Support Vector Machine Dan Naïve Bayes Dalam Klasifikasi Peluang Penyakit Serangan Jantung. *Indonesian Journal of Business Intelligence (IJUBI)*, 5(2), 87-91.
- Sari, L. M., Yuliano, A., & Almudriki, A. (2019). Hubungan pengetahuan dan sikap keluarga terhadap kemampuan deteksi dini serangan stroke iskemik akut pada penanganan pre hospital. *Jurnal Kesehatan Perintis*, 6(1), 74-80.
- Utomo, D. P., & Mesran, M. (2020). Analisis komparasi metode klasifikasi data mining dan reduksi atribut pada data set penyakit jantung. *Jurnal Media Informatika Budidarma*, 4(2), 437-444.