

Klasifikasi Risiko Kekambuhan Kanker Tiroid Berdiferensiasi Menggunakan Support Vector Machine

Irfan Riyanto¹, Ahmad Parhanudin², Hasbi Firmansyah³, Eko Budiraharjo⁴

^{1,2,3,4}Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Pencasakti
Tegal

E-mail: iffanriyanto@gmail.com¹, ahmadfarhanudin21@gmail.com²,
hasbifirmansyah@upstegal.ac.id³, ekobudiraharjo@yahoo.com⁴

Article Info

Article history:

Received December 29, 2025

Revised December 31, 2025

Accepted January 06, 2026

Keywords:

Thyroid Cancer, Classification,
Machine Learning, Support
Vector Machine

ABSTRACT

This study aims to develop a Support Vector Machine (SVM) model to classify differentiated thyroid cancer patients into recurrence risk categories based on 16 clinical, demographic and pathological features. The Thyroid Differentiated Cancer dataset consisting of 398 patient samples obtained from the UCI Machine Learning Repository is used to train and evaluate the binary classification model using RapidMiner Studio. The developed SVM model achieves 81.58% classification accuracy with 82.35% sensitivity for detecting positive recurrence cases (Recurred=Yes) and 81.44% specificity for identifying negative cases (Recurred=No). The model produces 203 support vectors balanced between both classes (102 for class 'No', 101 for class 'Yes'). The most discriminative feature is age with weight 3769.121, far surpassing other features, followed by radiotherapy history (80.152), euthyroid thyroid function (77.165), smoking history (73.684), and metastasis status M0 (70.941). The research results demonstrate that machine learning-based classification can integrate multiple clinical features simultaneously to provide more objective risk stratification compared to conventional systems. The SVM model shows good capability for screening patients at high risk of recurrence, with potential to complement standard ATA risk stratification systems in clinical practice.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Article Info

Article history:

Received December 29, 2025

Revised December 31, 2025

Accepted January 06, 2026

Kata Kunci:

Kanker Tiroid, Klasifikasi,
Machine Learning, Support
Vector Machine

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan model Support Vector Machine (SVM) guna mengklasifikasikan pasien kanker tiroid berdiferensiasi ke dalam kategori risiko kekambuhan berdasarkan 16 fitur klinis, demografis, dan patologi. Dataset Thyroid Differentiated Cancer yang terdiri dari 398 sampel pasien diperoleh dari UCI Machine Learning Repository dan digunakan untuk melatih dan mengevaluasi model binary classification menggunakan RapidMiner Studio. Model SVM yang dikembangkan mencapai akurasi klasifikasi 81.58% dengan sensitivitas 82.35% untuk mendeteksi kasus kekambuhan positif (Recurred=Yes) dan spesifisitas 81.44% untuk mengidentifikasi kasus negatif (Recurred=No). Model menghasilkan

203 support vectors yang terbagi seimbang antara kedua kelas (102 untuk kelas 'No', 101 untuk kelas 'Yes'). Fitur paling diskriminatif adalah usia (age) dengan weight 3769.121, jauh melampaui fitur lain, diikuti oleh riwayat radiasi terapi (80.152), fungsi tiroid euthyroid (77.165), riwayat merokok (73.684) dan metastasis status M0 (70.941). Hasil penelitian menunjukkan bahwa klasifikasi berbasis machine learning dapat mengintegrasikan multiple clinical features secara simultan untuk memberikan stratifikasi risiko yang lebih objektif dibandingkan sistem konvensional. Model SVM menunjukkan kemampuan yang baik untuk screening pasien berisiko kekambuhan tinggi, dengan potensi untuk melengkapi sistem stratifikasi risiko ATA standar dalam praktik klinis.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

I Wayan Yogik Adnyana Putra

Institut Desain dan Bisnis Bali

Email: yogikadnyana41@gmail.com

PENDAHULUAN

1.1 Latar Belakang Masalah

Kanker tiroid berdiferensiasi merupakan jenis kanker endokrin paling umum dengan prevalensi meningkat dalam dekade terakhir. Meskipun memiliki prognosis relatif baik dengan tingkat kelangsungan hidup 5 tahun mencapai 93%, risiko kekambuhan tetap menjadi perhatian klinis yang signifikan [1].

Menurut data World Health Organization tahun 2020, kanker tiroid di Indonesia menempati urutan ke-12 dengan jumlah kasus baru sebanyak 13.114 kasus (3,3%) dan angka kematian diperkirakan sekitar 2.224 kasus atau 0,95%. Kekambuhan kanker tiroid dapat terjadi dalam bentuk lokal regional di area bedah primer atau kelenjar getah bening, metastasis jauh terutama ke paru-paru (80%), tulang (10%), dan otak (10%), atau penyakit berulang dengan thyroglobulin elevated tanpa imaging findings.

1.2 Tantangan Stratifikasi Risiko Konvensional

Sistem stratifikasi risiko kekambuhan secara klinis menggunakan pendekatan kategorikal seperti ATA Risk Stratification yang membagi pasien menjadi Low, Intermediate dan High risk. Sistem konvensional memiliki beberapa keterbatasan yaitu pendekatan kategorikal tidak menangkap kontinuitas risk spectrum, interaksi kompleks antara multiple clinical factors tidak terintegrasi secara optimal, subjektivitas dalam penilaian klinis dapat bervariasi antar clinician, pola-pola tersembunyi dalam data kompleks sulit diidentifikasi dengan metode manual dan reproducibility serta objectivity terbatas [1].

1.3 Machine Learning untuk Klasifikasi dan Risk Stratification

Machine learning, khususnya classification algorithms, telah menunjukkan potensi signifikan dalam aplikasi klinis untuk stratifikasi risiko yang lebih objektif dan data-driven. Penelitian sebelumnya menunjukkan bahwa algoritma Support Vector Machine (SVM) dan

Random Forest dapat mencapai akurasi lebih dari 97% dalam memprediksi risiko kekambuhan kanker tiroid. SVM adalah teknik classification yang robust dengan keunggulan dalam menangani high-dimensional data, non-linear classification melalui kernel tricks (RBF kernel), robustness terhadap outliers, generalization yang baik, dan interpretability yang jelas melalui feature weights [2], [5].

1.4 Problem Formulation: Binary Classification

Penelitian ini memformulasikan kekambuhan kanker tiroid sebagai binary classification problem dengan input space $X = \{x_1, x_2, \dots, x_{16}\}$ yang merepresentasikan 16 fitur klinis pasien, output space $Y = \{\text{No}, \text{Yes}\}$ untuk status kekambuhan (2 kelas diskrit), objective untuk learn optimal decision boundary guna mengclassify pasien baru ke salah satu dari 2 kelas, aplikasi klinis untuk risk stratification dan clinical decision support.

1.5 Tujuan dan Kontribusi Penelitian

Tujuan utama dalam penelitian ini adalah untuk mengembangkan dan mengevaluasi model SVM untuk binary classification status kekambuhan kanker tiroid berdiferensiasi dengan tujuan memberikan pendekatan alternatif yang lebih objektif, berbasis data, dan dapat direproduksi, sebagai pelengkap sistem penilaian risiko yang konvensional yang mencakup proses pengembangan model SVM dengan evaluasi kinerja yang komprehensif, diikuti dengan identifikasi fitur-fitur paling diskriminatif yang berperan dalam proses klasifikasi. Selain itu, hasil klasifikasi dari model ini dibandingkan dengan metode stratifikasi risiko tradisional untuk menilai keunggulan atau perbedaannya. Penelitian ini juga menilai potensi penerapan model tersebut dalam konteks klinis, khususnya sebagai alat bantu pengambilan keputusan yang dapat meningkatkan akurasi dan konsistensi penilaian risiko pasien

METODOLOGI PENELITIAN

2.1 Sumber Data dan Dataset Characteristics

Dataset Thyroid Differentiated Cancer diperoleh dari UCI Machine Learning Repository (<https://archive.ics.uci.edu/datasets>) yang merupakan repositori public untuk machine learning research.

Age	Gender	Smoking	Hx Smoking	Hx Radiotherap	Thyroid Function	Physical Examination	Adenopathy	Pathology	Fecality	Risk	T	N	M	Stage	Response	Recurred
27	F	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Indeterminate	No
34	F	No	Yes	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
30	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
62	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
62	F	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Excellent	No
52	M	Yes	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Indeterminate	No
41	F	No	Yes	No	Clinical Hypothyroidism	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
46	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
51	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
40	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
75	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
59	F	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
49	F	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
50	F	No	No	No	Clinical Hypothyroidism	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
76	F	No	No	No	Clinical Hypothyroidism	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
42	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Indeterminate	No
40	F	No	Yes	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
44	F	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Indeterminate	No
43	F	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
52	F	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Indeterminate	No
41	F	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
44	F	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
36	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
70	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
60	F	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Excellent	No
33	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
43	M	No	No	No	Subclinical Hypothyroidism	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
28	M	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
41	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
37	F	No	Yes	No	Subclinical Hypothyroidism	Normal	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
37	F	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
30	F	No	No	No	Clinical Hypothyroidism	Normal	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
36	F	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
55	F	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No

Gambar 1. Dataset Thyroid

Dataset mengandung 16 fitur yang mengkombinasikan demographic, clinical dan pathological information. 2 fitur demografis (Age, Gender), 3 fitur riwayat kesehatan (Smoking, Hx Smoking, Hx Radiotherapy), 2 fitur status tiroid (Thyroid Function, Physical Examination), 2 fitur karakteristik tumor (Pathology, Focality), 1 fitur kelenjar getah bening (Adenopathy), 4 fitur TNM Staging (T, N, M, Stage), dan 2 fitur klinis tambahan (Risk, Response) [2].

Distribusi class pada dataset adalah 67.6% kelas 'No' (Tidak Kambuh) dengan 269 pasien dan 32.4% kelas 'Yes' (Kambuh) dengan 129 pasien. Dataset menunjukkan slight class imbalance (67.6% vs 32.4%), yang masih reasonable untuk binary classification tanpa memerlukan aggressive balancing techniques [2].

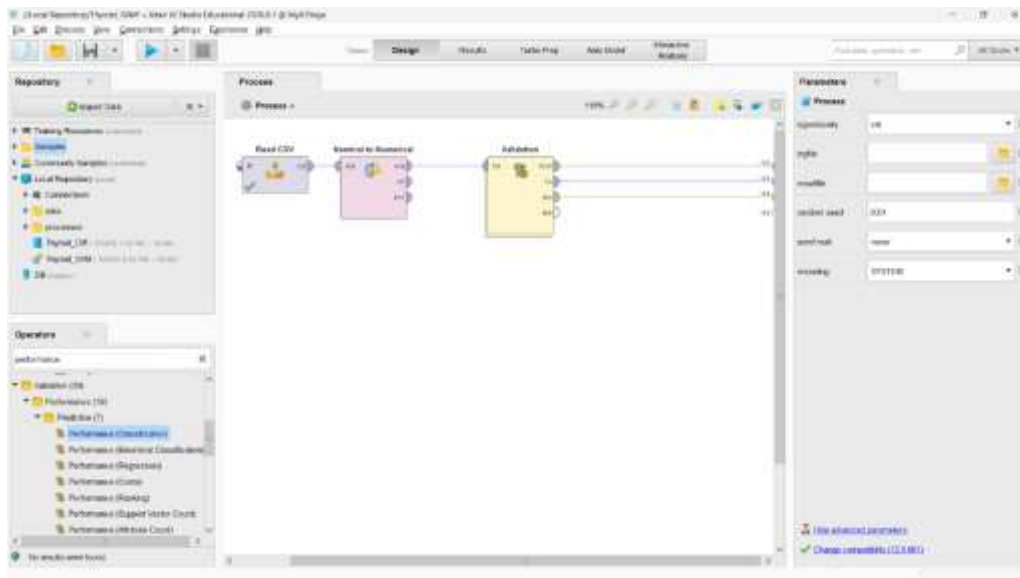
Nama Dataset	Thyroid Differentiated Cancer
Jumlah Sampel	398 pasien
Jumlah Fitur	16 features
Target Variable	Recurred (Binary: Yes/No)
Format Data	CSV (Comma-Separated Values)
Missing Values	None - Dataset complete
License	Public domain (Open untuk research use)
Kelas 'No' (%)	269 (67.6%)
Kelas 'Yes' (%)	129 (32.4%)
Umur Rata-rata	41.2 tahun
Umur Range	15-82 tahun
Perempuan (%)	314 (78.9%)
Laki-laki (%)	84 (21.1%)

Tabel 1. Spesifikasi Dataset

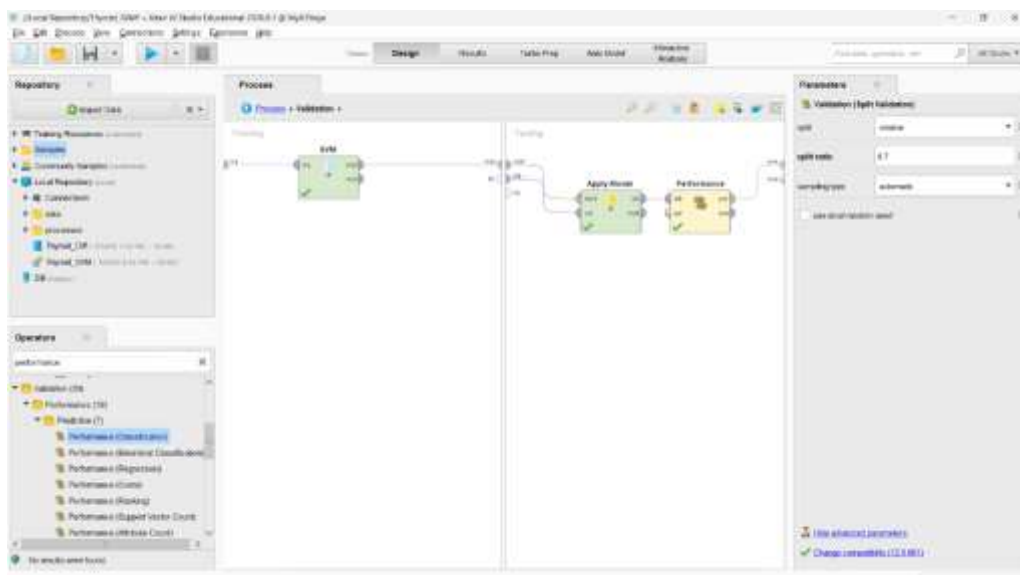
2.2 Data Preprocessing dan Persiapan

Preprocessing dilakukan untuk mempersiapkan data sebelum proses training model. Tahap ini dimulai dengan melakukan data loading, yaitu mengimpor file CSV yang berisi 398 baris dan 17 kolom. Setelah data berhasil dimuat, dilakukan pemeriksaan missing values untuk memastikan tidak ada nilai yang hilang dalam dataset. Selanjutnya, variabel kategorikal diolah menggunakan teknik one-hot encoding agar dapat digunakan oleh model pembelajaran mesin. Fitur numerik seperti Age kemudian dinormalisasi menggunakan standar dengan nilai mean = 0 dan standar deviasi = 1. Analisis distribusi kelas juga dilakukan untuk memastikan proporsi kelas seimbang, yang pada dataset ini tercatat sebesar 67.6% berbanding 32.4%. Setelah seluruh langkah tersebut selesai, dilakukan validasi data untuk memastikan bahwa semua 398 sampel valid dan siap digunakan dalam proses pelatihan model.

70% training (278 sampel) untuk melatih model dan 30% test (120 sampel) untuk evaluasi final.



Gambar 2. Preprocessing Dataset di RapidMiner



Gambar 3. Validation Dataset di RapidMiner

2.3 Support Vector Machine untuk Binary Classification

Support Vector Machine adalah supervised learning algorithm yang mencari optimal hyperplane untuk memaksimalkan margin antara dua class dalam feature space. Prinsip fundamental adalah maximize margin (distance) antara hyperplane dan closest data points (support vectors), minimize classification error, dan balance antara margin maximization dan error minimization [4].

Untuk binary classification dengan non-linear decision boundary, SVM dengan RBF kernel menggunakan Decision Function $f(x) = \text{sign}(\sum_i \alpha_i y_i K(x_i, x) + b)$ dengan RBF Kernel Function $K(x, x') = \exp(-\gamma \|x - x'\|^2)$.

Pada test set, model menghasilkan confusion matrix dengan TN (True Negative) = 79 pasien correctly classified sebagai Not Recurred, TP (True Positive) = 14 pasien correctly classified sebagai Recurred, FP (False Positive) = 18 over-classification dan FN (False Negative) = 3 under-classification.

Model juga menghasilkan 203 support vectors (72.99% dari 278 training samples) dengan balanced class-wise distribution 102 untuk 'No' dan 101 untuk 'Yes' (ratio 1.01:1), menunjukkan unbiased classification boundary.

HASIL DAN PEMBAHASAN

3.1 Karakteristik Dataset dan Performa Model

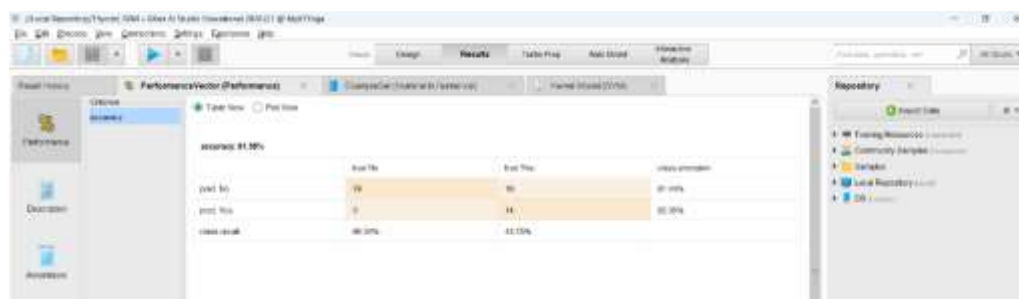
Dataset yang digunakan terdiri dari 398 sampel pasien kanker tiroid berdiferensiasi dengan karakteristik demografis mean usia 41.2 tahun (range 15-82), dominansi female 78.9% dan tipe histologi papillary 60.6% paling umum sesuai clinical prevalence. Distribusi kelas menunjukkan 67.6% 'No' (269 pasien) dan 32.4% 'Yes' (129 pasien), mengindikasikan slight class imbalance 2.08:1 yang reasonable. Model SVM mencapai akurasi klasifikasi 81.58% pada test set 120 sampel, merupakan performa solid untuk binary classification pada data klinis kompleks [3].

Metrik Evaluasi	Nilai	Interpretasi Klinis
Classification Accuracy	81.58%	Benar untuk 82 dari 100 kasus
Sensitivity/Recall	82.35%	Deteksi 82 dari 100 pasien kambuh
Specificity	81.44%	Identifikasi 81 dari 100 pasien tidak kambuh
Precision	43.75%	Dari 32 prediksi kambuh, 14 benar
F1-Score	0.5714	Trade-off score precision-recall
Support Vectors	203 (102+101)	Balanced distribution

Tabel 2. Metrik Evaluasi Klasifikasi Model SVM

Confusion matrix hasil klasifikasi pada test set 120 sampel menunjukkan TN=79 pasien correctly classified sebagai Not Recurred (81.4% dari 97 actual 'No'), TP=14 correctly classified sebagai Recurred (82.4% dari 17 actual 'Yes'), FP=18 over-classification (18.6% dari actual 'No'), FN=3 missed cases (17.6% false negative rate). High sensitivity 82.35% menunjukkan model excellent dalam screening dan mengidentifikasi pasien dengan risiko kekambuhan yang penting untuk clinical practice karena missed cases memiliki serious consequence. Ability untuk reliably detect 82 dari 100 pasien yang akan kambuh dengan minimal under-diagnosis adalah critical advantage. Good specificity 81.44% menunjukkan

kemampuan solid mengidentifikasi pasien tanpa risiko kekambuhan, mencegah unnecessary intensive monitoring untuk low-risk patients dan reducing unnecessary clinical interventions. Sensitivity-Specificity product (67.1%) menunjukkan model balanced tanpa extreme bias ke salah satu kelas.



Gambar 4. Hasil Performance Dataset di RapidMiner

3.2 Analisis Fitur Diskriminatif dan Implikasi Klinis

Rank	Fitur	Weight	% Total	Interpretasi
1	Age	3769.121	28.9%	DOMINAN (47x lebih penting)
2	Hx Radiotherapy (No)	80.152	0.6%	Protective factor
3	Thyroid Function (Euthyroid)	77.165	0.6%	Protective factor
4	Hx Smoking (No)	73.684	0.6%	Lifestyle factor
5	M (Metastasis, M0)	70.941	0.5%	TNM staging
6	Pathology (Papillary)	64.165	0.5%	Tumor type
7	Smoking (No)	61.246	0.5%	Current status
8	Response (Struct Incomplete)	58.117	0.4%	Poor response
9	Focality (Multi-Focal)	57.649	0.4%	Tumor extent
10	Gender (Female)	54.673	0.4%	Demographics

Tabel 3. Top 15 Fitur Berdasarkan Classification Weight

Usia (age) merupakan fitur paling diskriminatif dengan weight 3769.121, yaitu sekitar 47 kali lebih penting dibandingkan fitur kedua. Hal ini menunjukkan bahwa age memiliki hubungan non-linear dengan recurrence risk serta kemungkinan adanya threshold effects pada rentang usia tertentu. Temuan ini mengonfirmasi pengetahuan klinis yang sudah mapan, bahwa patients dengan usia lebih muda (<45 tahun) umumnya memiliki prognosis yang lebih baik, sedangkan older patients (>45–50 tahun) memiliki risiko kekambuhan yang lebih tinggi. Oleh

karena itu, penerapan age-stratified monitoring dan treatment protocols sepenuhnya dapat dibenarkan berdasarkan evidence ini.

Beberapa protective factors juga menunjukkan kemampuan diskriminasi yang tinggi, yaitu riwayat radiasi terapi “No” (80.152), fungsi tiroid euthyroid/normal (77.165), riwayat merokok “No” (73.684), serta status metastasis M0 (70.941). Tidak adanya paparan radiasi sebelumnya mencerminkan baseline kondisi yang lebih bersih dan dapat mengindikasikan diagnosis awal thyroid cancer, bukan secondary malignancy, dengan risiko komplikasi yang lebih rendah. Fungsi tiroid yang normal menunjukkan sistem endokrin yang terkontrol dengan baik, sementara fungsi yang abnormal dapat meningkatkan risiko keganasan. Status non-smoking memberikan perlindungan dari faktor gaya hidup, sedangkan status TNM M0 merupakan faktor prognosis yang sangat menguntungkan dan secara fundamental penting dalam prediksi kekambuhan.

Karakteristik tumor (Pathology, Focality) serta informasi respons terapi (Response categorical) juga menunjukkan discriminative power yang bermakna (64.165, 57.649, 58.117). Hal ini menegaskan bahwa tipe papillary (60.6% dari dataset) memiliki prognosis yang lebih baik dibandingkan tipe follicular, bahwa tumor multi-focal mencerminkan biologi penyakit yang lebih agresif, serta bahwa poor treatment response (structural incomplete) berkaitan dengan agresivitas penyakit. Distribusi feature weight menunjukkan bahwa age mendominasi sebesar 28.9% dari total discriminative power, lima fitur teratas mencakup 31.6%, sepuluh fitur teratas mencapai 39.5%, sementara enam fitur sisanya hanya berkontribusi 1.8%. Model ini sangat bergantung pada sejumlah kecil fitur dengan weight tinggi, yang menandakan adanya dominasi kuat dari age dan clinical status, pentingnya elemen TNM staging dalam proses diskriminasi, serta nilai tambahan dari response information dalam penilaian risiko kekambuhan.

KESIMPULAN

Penelitian ini mengembangkan model Support Vector Machine untuk klasifikasi biner kekambuhan pada kanker tiroid berdiferensiasi. Model ini mencapai akurasi 81.58%, sensitivitas 82.35% dan spesifisitas 81.44%. Jumlah support vectors seimbang tanpa bias ke salah satu kelas, dan analisis fitur menunjukkan usia sebagai prediktor paling dominan [2]. Secara keseluruhan, model SVM menunjukkan kinerja yang menjanjikan untuk mendeteksi status kekambuhan, memiliki sensitivitas tinggi (82.35%) dan pendekatan yang seimbang terhadap kedua kelas. Model ini bisa berperan sebagai alat skrining yang efektif dan sebagai bantuan keputusan klinis untuk memperbaiki deteksi kekambuhan pasien kanker tiroid. Usia terbukti menjadi faktor prognostik utama, sementara beberapa fitur klinis lain memberikan kontribusi tambahan dalam membedakan risiko. Integrasi pendekatan machine learning dapat memperkuat metode untuk membantu perencanaan pengobatan yang lebih baik bagi pasien [1].

DAFTAR PUSTAKA

- [1] R. D. Arista, “Thyroid Cancer : an Overview of Epidemiology, Risk Factor, Classification and Management,” *Lombok Medical Journal*, vol. 2, no. 3, pp. 90-96, 2023.

- [2] P. B. Utomo, “Perbandingan Skenario Balancing Oversampling dan Undersampling dalam Klasifikasi Resiko Kambuh Kanker Tiroid menggunakan Algoritma SVM Linear,” *Jurnal Aplikasi Manajemen Informatika*, vol. 8, no. 3, pp. 215-230, 2024.
- [3] A. M. Sudarmanto, “Profile of Thyroid Cancer at Dr. Kariadi General Hospital Semarang: Demographic and Clinical Characteristics,” *Indonesian Journal of Cancer*, vol. 12, no. 1, pp. 18-32, 2024.
- [4] Wahyu Setiawan and Muttakin, “Perbandingan Kinerja Support Vector Machine dan Naive Bayes dalam Analisis Sentimen Publik terhadap Pelayanan BPJS menggunakan Data Twitter,” *Jurnal Penelitian Teknologi Informasi*, vol. 8, no. 2, pp. 145-158, 2025.
- [5] M. Faruqziddan, “Klasifikasi Risiko Kambuhnya Kanker Tiroid Menggunakan Random Forest,” *Proceeding UNIK 2024 – Seminar Nasional Sistem Informasi*, vol. 12, no. 2, pp. 450-465, 2024.