

Optimasi Performa Model SVM dan Random Forest untuk Klasifikasi Kanker Payudara Menggunakan Penyetelan Hyperparameter

Adil Setiawan¹, Andri Armaginda Siregar^{2*}, Nanda Setiawan³, Jalaluddin Nasution⁴, Naufal Dhiya Putra Dalimunthe⁵, Farhan Sardy Abdillah⁶

^{1,2,3,4,5,6}Faculty of Engineering & Computer Science, master of computer science, Universitas Potensi Utama, Medan, Indonesia

Email: ¹Adio165@gmail.com, ^{2*}andriarmagindasiregar@gmail.com, ³nandasetiawansn27@gmail.com,

⁴jalaluddin.nasution@uinsu.ac.id, ⁵naufaldhiyaputradalimunthe@gmail.com, ⁶fsabdillah19@gmail.com

(*Email Corresponding Author: andriarmagindasiregar@gmail.com)

Received: January 10, 2026 | Revision: January 21, 2026 | Accepted: January 23, 2026

Abstrak

Kanker payudara merupakan salah satu penyakit dengan tingkat morbiditas dan mortalitas yang tinggi, sehingga deteksi dini dan diagnosis yang akurat menjadi faktor kunci dalam meningkatkan keberhasilan terapi. Perkembangan *machine learning* memberikan peluang besar dalam mendukung proses klasifikasi kanker payudara berbasis data medis. Penelitian ini bertujuan untuk mengevaluasi dan mengoptimalkan kinerja algoritma *Support Vector Machine* (SVM) dan *Random Forest* dalam klasifikasi kanker payudara melalui penerapan *Hyperparameter tuning*. Dataset yang digunakan adalah *Breast Cancer Wisconsin (Diagnostic)*, yang terdiri dari 569 sampel hasil pemeriksaan *Fine Needle Aspirate* (FNA) dengan 30 fitur numerik serta dua kelas target, yaitu benign dan malignant. Metodologi penelitian meliputi tahap pra-pemrosesan data, pembagian data menggunakan stratified train-test split, pelatihan model *baseline*, optimasi *Hyperparameter* menggunakan *GridSearchCV* dengan 5-fold cross-validation, serta evaluasi performa model menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, *confusion matrix*, dan *ROC-AUC*. Hasil eksperimen menunjukkan bahwa kedua model mampu mencapai performa klasifikasi yang tinggi dengan tingkat akurasi sekitar 97% pada data uji. *Hyperparameter tuning* memberikan peningkatan performa yang lebih konsisten pada model SVM, khususnya pada *recall* kelas malignant dan nilai *AUC*, sedangkan *Random Forest* menunjukkan performa yang relatif stabil sebelum dan sesudah optimasi. Temuan ini menegaskan bahwa *Hyperparameter tuning* berperan penting dalam validasi dan stabilitas model, meskipun tidak selalu menghasilkan peningkatan performa numerik yang signifikan pada *dataset benchmark*.

Kata Kunci: Kanker payudara, *Machine learning*, *Support Vector Machine*, *Random Forest*, *Hyperparameter tuning*, Klasifikasi medis

Abstract

Breast cancer is one of the diseases with high morbidity and mortality rates; therefore, early detection and accurate diagnosis are key factors in improving treatment success. Advances in machine learning offer significant opportunities to support breast cancer classification based on medical data. This study aims to evaluate and optimize the performance of the Support Vector Machine (SVM) and Random Forest algorithms in breast cancer classification through the application of hyperparameter tuning. The dataset used is the Breast Cancer Wisconsin (Diagnostic) dataset, which consists of 569 samples obtained from Fine Needle Aspirate (FNA) examinations, with 30 numerical features and two target classes, namely benign and malignant. The research methodology includes data preprocessing, data splitting using stratified train-test split, baseline model training, hyperparameter optimization using GridSearchCV with 5-fold cross-validation, and model performance evaluation using accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC metrics. The experimental results show that both models are able to achieve high classification performance, with accuracy levels of approximately 97% on the test data. Hyperparameter tuning provides a more consistent performance improvement for the SVM model, particularly in terms of malignant-class recall and AUC values, while Random Forest demonstrates relatively stable performance before and after optimization. These findings confirm that hyperparameter tuning plays an important role in model validation and stability, although it does not always result in significant numerical performance improvements on benchmark datasets.

Keywords: *Breast cancer, Machine learning, Support Vector Machine, Random Forest, Hyperparameter tuning, Medical classification*

1. PENDAHULUAN

Kanker merupakan salah satu permasalahan kesehatan global yang hingga saat ini masih menjadi penyebab utama tingginya angka kematian di berbagai negara. Di antara berbagai jenis kanker, kanker payudara menempati posisi teratas sebagai kanker yang paling banyak diderita oleh perempuan, dengan tingkat kejadian dan mortalitas yang terus meningkat setiap tahunnya. Tantangan utama dalam penanganan kanker payudara terletak pada proses diagnosis, khususnya pada tahap awal penyakit, di mana gejala klinis sering kali belum terlihat secara jelas [1].

Perkembangan pesat kecerdasan buatan (*Artificial Intelligence/AI*) telah membawa perubahan signifikan dalam bidang medis, terutama dalam pemanfaatan *machine learning* dan *deep learning* untuk analisis data kesehatan. Berbagai penelitian menunjukkan bahwa pendekatan berbasis AI mampu mengolah data medis dalam jumlah besar dan kompleks, serta mengekstraksi pola-pola tersembunyi yang sulit diidentifikasi melalui metode konvensional [2]. Dalam konteks onkologi, AI telah diterapkan pada berbagai jenis data, mulai dari citra medis seperti *mammogram*, CT scan, dan MRI, hingga data klinis dan molekuler untuk mendukung diagnosis, prediksi prognosis, serta pengambilan keputusan klinis [3], [4], [5]. Hasil-hasil penelitian tersebut mengindikasikan bahwa pemanfaatan AI berpotensi meningkatkan akurasi dan efisiensi sistem diagnosis kanker.

Meskipun *deep learning* banyak digunakan pada analisis citra medis, algoritma *machine learning* klasik seperti *Support Vector Machine* (SVM) dan *Random Forest* (RF) tetap memiliki peran penting, khususnya pada dataset tabular dengan ukuran menengah dan fitur numerik yang terstruktur [6]. SVM dikenal memiliki kemampuan yang baik dalam membangun batas keputusan non-linear melalui penggunaan kernel, sehingga efektif untuk memisahkan kelas pada data berdimensi tinggi. Di sisi lain, *Random Forest* sebagai metode *ensemble learning* mampu meningkatkan stabilitas dan ketahanan model terhadap *noise* dengan mengkombinasikan banyak pohon keputusan. Kedua algoritma ini telah digunakan secara luas dalam penelitian kanker, termasuk untuk klasifikasi kanker payudara, dan menunjukkan performa yang kompetitif [7], [8].

Namun demikian, performa algoritma *machine learning* tidak hanya ditentukan oleh pemilihan model, tetapi juga sangat bergantung pada konfigurasi *Hyperparameter* yang digunakan. *Hyperparameter* yang tidak optimal dapat menyebabkan model mengalami *overfitting* atau *underfitting*, yang pada akhirnya berdampak pada penurunan kemampuan generalisasi model terhadap data baru. Dalam konteks diagnosis medis, kesalahan klasifikasi memiliki implikasi yang serius, terutama kesalahan *false negative*, dimana pasien dengan kanker diklasifikasikan sebagai non-kanker. Kesalahan jenis ini berpotensi menunda penanganan medis dan meningkatkan risiko klinis bagi pasien. Oleh karena itu, optimasi *Hyperparameter* menjadi langkah penting untuk memastikan bahwa model tidak hanya memiliki akurasi tinggi, tetapi juga sensitivitas yang memadai dalam mendeteksi kasus kanker.

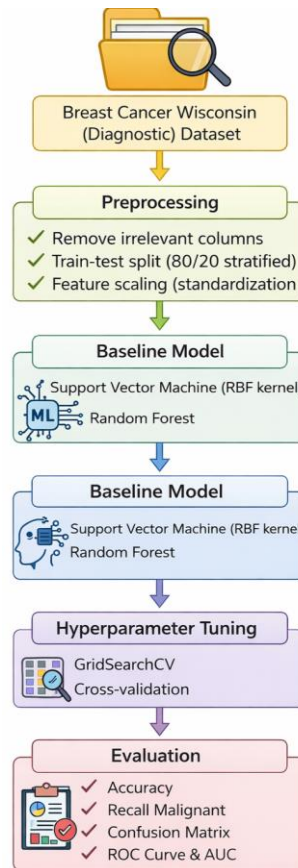
Berdasarkan tinjauan terhadap penelitian terdahulu, sebagian besar studi lebih menitikberatkan pada penggunaan algoritma tertentu atau perbandingan performa antar model secara umum, tanpa melakukan analisis mendalam terhadap dampak *Hyperparameter tuning* terhadap distribusi kesalahan klasifikasi. Beberapa penelitian melaporkan akurasi yang tinggi, namun tidak selalu menyertakan analisis komprehensif menggunakan metrik evaluasi yang berorientasi pada keselamatan pasien, seperti *recall* kelas malignant dan *area under the ROC curve* (ROC-AUC) [1], [2], [5]. Selain itu, masih terbatas penelitian yang secara eksplisit membandingkan performa model *baseline* dan model teroptimasi pada *dataset benchmark* kanker payudara dengan pendekatan eksperimen yang sistematis dan reproduibel. Kondisi ini menunjukkan adanya celah penelitian terkait evaluasi dan optimasi model *machine learning* klasik untuk diagnosis kanker payudara berbasis data tabular.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengoptimalkan dan mengevaluasi performa model *Support Vector Machine* dan *Random Forest* dalam klasifikasi kanker payudara menggunakan Dataset *Breast Cancer Wisconsin (Diagnostic)*. Optimasi model dilakukan melalui pendekatan *Hyperparameter tuning* berbasis *GridSearchCV*, dengan fokus tidak hanya pada peningkatan akurasi, tetapi juga pada sensitivitas deteksi kanker dan kemampuan diskriminasi model [9]. Untuk menjawab celah penelitian yang ada, kontribusi utama penelitian ini dirumuskan sebagai berikut: (1) menyajikan evaluasi komprehensif terhadap performa model *baseline* dan model teroptimasi menggunakan berbagai metrik evaluasi, termasuk akurasi, *recall* kelas malignant, *confusion matrix*, dan ROC-AUC; (2) menganalisis secara empiris pengaruh *Hyperparameter tuning* terhadap sensitivitas deteksi kanker payudara, dengan penekanan pada pengurangan jumlah *false negative* yang memiliki dampak klinis tinggi; (3) melakukan perbandingan kinerja yang sistematis antara algoritma SVM dan *Random Forest* untuk mengidentifikasi model yang paling sesuai pada dataset *Breast Cancer Wisconsin (Diagnostic)*; dan (4) menyediakan kerangka eksperimen yang reproduibel dan terdokumentasi dengan baik sebagai referensi bagi penelitian selanjutnya dan pengembangan sistem pendukung diagnosis kanker berbasis *machine learning*.

Secara keseluruhan, penelitian ini diharapkan dapat memberikan kontribusi ilmiah dalam bidang penerapan *machine learning* untuk diagnosis kanker payudara, khususnya dalam menunjukkan pentingnya optimasi *Hyperparameter* dan pemilihan metrik evaluasi yang tepat. Dengan menekankan aspek sensitivitas dan reliabilitas model, hasil penelitian ini diharapkan dapat mendukung pengembangan sistem diagnosis yang lebih akurat, aman, dan aplikatif dalam praktik klinis, serta menjadi dasar bagi penelitian lanjutan di bidang kecerdasan buatan medis.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk mengevaluasi performa algoritma *machine learning* dalam klasifikasi kanker payudara. Metodologi penelitian dirancang secara sistematis agar setiap tahapan dapat direplikasi dan memberikan hasil yang objektif. Fokus utama penelitian ini adalah menganalisis pengaruh *Hyperparameter tuning* terhadap performa model *Support Vector Machine* (SVM) dan *Random Forest* (RF) pada dataset kanker payudara berbasis data tabular.



Gambar 1. Diagram Alur Metodologi Penelitian

Secara umum, alur metodologi penelitian terdiri dari lima tahapan utama, yaitu akuisisi data, pra-pemrosesan data, pemodelan *machine learning*, optimasi *Hyperparameter*, dan evaluasi performa model. Alur konseptual penelitian ini dapat direpresentasikan dalam bentuk diagram alir metodologi, sebagaimana ditunjukkan pada Gambar 1, yang menggambarkan hubungan antar tahapan penelitian dari awal hingga akhir.

2.1 Dataset dan Akuisisi Data

Dataset yang digunakan dalam penelitian ini adalah *Breast Cancer Wisconsin (Diagnostic)* dataset, yang merupakan *dataset benchmark* dalam penelitian klasifikasi kanker payudara. Dataset ini terdiri dari 569 sampel hasil pemeriksaan *Fine Needle Asperate (FNA)* terhadap jaringan payudara, dengan dua kelas target, yaitu jaringan jinak (Benign) dan jaringan ganas (Malignant). Setiap sampel direpresentasikan oleh 30 fitur numerik yang menggambarkan karakteristik morfologi inti sel, termasuk ukuran, tekstur, bentuk, dan kompleksitas permukaan sel.

Tabel 1. Deskripsi Dataset *Breast Cancer Wisconsin (Diagnostic)*

Parameter	Keterangan
Jumlah Sampel	569
Jumlah Fitur	30
Kelas	Benign, Malignant
Sumber Data	<i>Fine Needle Asperate (FNA)</i>
Jenis Data	Numerik (Tabular)

Penggunaan *dataset benchmark* dalam penelitian klasifikasi medis bertujuan untuk memastikan keterbandingan hasil eksperimen dengan penelitian terdahulu serta meningkatkan reproduibilitas penelitian [2], [5], [10] Dataset ini dipilih karena memiliki struktur data yang bersih, tidak mengandung nilai hilang pada fitur utama, serta memiliki distribusi kelas yang relatif seimbang, sehingga sesuai untuk evaluasi performa algoritma *machine learning* klasik.

2.2 Pra-pemrosesan Data

Pra-pemrosesan data dilakukan untuk memastikan kualitas data dan mengurangi potensi bias sebelum proses pelatihan model. Pada tahap ini, atribut yang tidak relevan terhadap proses klasifikasi, seperti kolom identifikasi dan

kolom kosong, dihapus dari dataset. Selanjutnya, label diagnosis yang semula berbentuk kategorikal dikonversi ke dalam bentuk numerik agar dapat diproses oleh algoritma *machine learning*.

Dataset kemudian dipisahkan menjadi fitur (X) dan label target (y), serta dibagi menjadi data latih dan data uji dengan proporsi 80% dan 20% [6]. Proses pembagian dilakukan menggunakan metode stratified sampling untuk mempertahankan distribusi kelas pada kedua subset, sebagaimana direkomendasikan dalam penelitian klasifikasi medis berbasis data tabular [2], [11].

Selain itu, dilakukan proses standarisasi fitur menggunakan metode z-score normalization, khususnya untuk mendukung kinerja algoritma yang sensitif terhadap skala fitur, seperti *Support Vector Machine*. Standarisasi bertujuan untuk memastikan bahwa setiap fitur memiliki kontribusi yang seimbang dalam proses pembelajaran model [10].

2.3 Pemodelan dan Eksperimen *Baseline*

Pada tahap pemodelan, dua algoritma *machine learning* digunakan, yaitu *Support Vector Machine* dan *Random Forest*. SVM dipilih karena kemampuannya dalam membangun batas keputusan non-linear melalui fungsi kernel, sehingga efektif untuk data berdimensi tinggi. Dalam penelitian ini, kernel Radial Basis Function (RBF) digunakan karena fleksibilitasnya dalam menangani hubungan non-linear antar fitur [7], [8].

Random Forest digunakan sebagai metode pembandingan karena merupakan algoritma *ensemble learning* yang mengkombinasikan banyak pohon keputusan untuk meningkatkan stabilitas dan ketahanan model terhadap *noise*. *Random Forest* juga relatif tidak sensitif terhadap overfitting dan mampu menangani interaksi antar fitur secara otomatis [10], [8].

Pada tahap awal, kedua model dilatih menggunakan konfigurasi parameter default untuk memperoleh model *baseline*. Model *baseline* ini berfungsi sebagai acuan awal sebelum dilakukan proses optimasi *Hyperparameter*.

2.4 *Hyperparameter Tuning*

Untuk meningkatkan performa dan stabilitas model, dilakukan proses *Hyperparameter tuning* menggunakan metode GridSearchCV. Pendekatan ini memungkinkan evaluasi berbagai kombinasi *Hyperparameter* secara sistematis dengan memanfaatkan validasi silang (k-fold cross-validation), sehingga mengurangi risiko overfitting dan meningkatkan kemampuan generalisasi model [2], [10].

Pada model SVM, *Hyperparameter* yang dioptimalkan meliputi parameter C sebagai pengendali regularisasi dan parameter gamma pada kernel Radial Basis Function (RBF) yang memengaruhi kompleksitas batas keputusan. Sementara itu, pada model *Random Forest*, *Hyperparameter* yang dioptimalkan mencakup jumlah pohon (*n_estimators*), kedalaman maksimum pohon (*max_depth*), serta parameter pemisahan node (*min_samples_split* dan *min_samples_leaf*).

Pendekatan *Hyperparameter tuning* dianggap penting dalam konteks aplikasi medis, karena konfigurasi parameter yang tidak optimal dapat menghasilkan kesalahan klasifikasi yang berdampak pada keselamatan pasien [1], [5].

2.5 Evaluasi Performa Model

Evaluasi performa model dilakukan untuk menilai kemampuan model dalam mengklasifikasikan data kanker payudara secara akurat dan konsisten. Proses evaluasi dilakukan menggunakan data uji yang tidak terlibat dalam proses pelatihan, sehingga memberikan gambaran objektif mengenai kemampuan generalisasi model [12].

Beberapa metrik evaluasi digunakan untuk memberikan penilaian yang komprehensif terhadap performa model. Metrik akurasi digunakan untuk mengukur proporsi prediksi yang benar secara keseluruhan. Namun, mengingat konteks diagnosis medis, penelitian ini juga menekankan penggunaan metrik *recall* untuk kelas malignant, yang mencerminkan kemampuan model dalam mendeteksi kasus kanker secara benar [1], [5].

Selain itu, *confusion matrix* digunakan untuk menyajikan distribusi hasil prediksi dalam bentuk *true positive*, *true negative*, *false positive*, dan *false negative* [13]. Untuk mengevaluasi kemampuan diskriminasi model secara keseluruhan, digunakan kurva *Receiver Operating Characteristic* (ROC) dan *Area Under the Curve* (AUC), yang menggambarkan hubungan antara tingkat deteksi benar dan tingkat kesalahan pada berbagai ambang keputusan [14].

Visualisasi evaluasi model dapat direpresentasikan dalam bentuk *confusion matrix* dan kurva ROC sebagai ilustrasi mekanisme evaluasi, sedangkan ringkasan metrik evaluasi disajikan dalam bentuk tabel perbandingan performa model [2], [10].

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil eksperimen serta pembahasan terkait performa model *Support Vector Machine* (SVM) dan *Random Forest* (RF) dalam klasifikasi kanker payudara menggunakan Dataset *Breast Cancer Wisconsin (Diagnostic)*. Seluruh hasil diperoleh dari evaluasi model pada data uji yang tidak digunakan selama proses pelatihan, sehingga mencerminkan kemampuan generalisasi model. Pembahasan difokuskan pada analisis performa model

baseline dan model teroptimasi, distribusi kesalahan klasifikasi, serta implikasi metodologis dan klinis dari hasil yang diperoleh.

3.1 Hasil Eksperimen Model *Baseline*

Eksperimen awal dilakukan dengan melatih model SVM dan *Random Forest* menggunakan konfigurasi parameter default sebagai model *baseline*. Tujuan tahap ini adalah memperoleh gambaran performa awal model tanpa intervensi optimasi *Hyperparameter*. Hasil evaluasi menunjukkan bahwa kedua algoritma mampu mencapai tingkat akurasi yang relatif tinggi pada data uji. Hal ini mengindikasikan bahwa Dataset *Breast Cancer Wisconsin (Diagnostic)* memiliki karakteristik fitur yang cukup informatif untuk dipelajari oleh algoritma *machine learning* klasik.

Meskipun demikian, evaluasi performa tidak hanya dilihat dari akurasi semata. Analisis terhadap *recall* kelas malignant menunjukkan bahwa model *baseline* masih menghasilkan sejumlah kesalahan klasifikasi pada kasus kanker payudara. Keberadaan *false negative* pada tahap *baseline* menjadi perhatian utama, mengingat kesalahan ini berpotensi berdampak langsung pada keselamatan pasien. Temuan ini menegaskan bahwa akurasi tinggi tidak selalu menjamin keandalan model dalam konteks diagnosis medis, serta memperkuat urgensi dilakukannya optimasi model.

Tabel 2. Perbandingan performa model *baseline* dan model teroptimasi

Model	Accuracy	Recall_Malignant	AUC
SVM (<i>Baseline</i>)	0.973684	0.928571	0.994709
SVM (Optimized)	0.982456	0.952381	0.996032
RF (<i>Baseline</i>)	0.973684	0.928571	0.992890
RF (Optimized)	0.973684	0.928571	0.992890

Hasil evaluasi model *baseline* dapat diringkas dalam Tabel 1, yang menyajikan metrik utama seperti accuracy, *recall* kelas malignant, dan AUC. Tabel ini berfungsi sebagai acuan awal untuk membandingkan dampak *Hyperparameter tuning* pada tahap selanjutnya.

3.2 Pengaruh *Hyperparameter Tuning* terhadap Performa Model

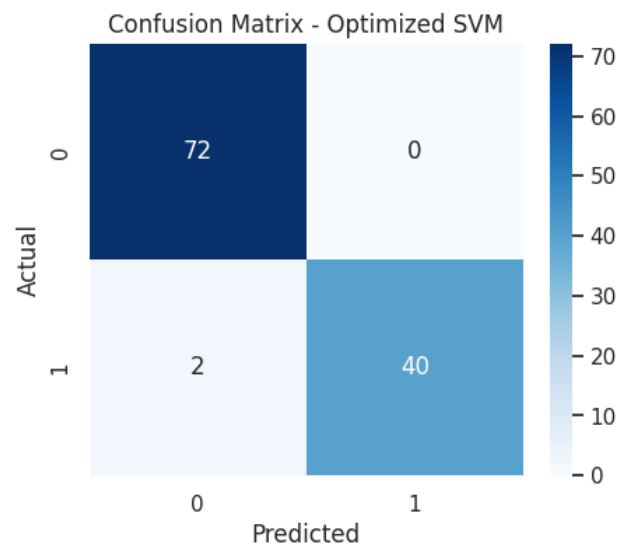
Hyperparameter tuning dilakukan menggunakan pendekatan GridSearchCV dengan validasi silang k-fold untuk memperoleh konfigurasi parameter terbaik bagi masing-masing model. Proses ini bertujuan untuk meningkatkan kemampuan model dalam menangkap pola data secara optimal dan mengurangi kesalahan klasifikasi.

Hasil eksperimen menunjukkan bahwa *Hyperparameter tuning* memberikan dampak yang berbeda pada kedua algoritma. Pada model SVM, proses optimasi menghasilkan peningkatan performa yang konsisten pada berbagai metrik evaluasi. Peningkatan ini mencerminkan kemampuan SVM dalam memanfaatkan penyesuaian parameter kernel untuk membangun batas keputusan yang lebih sesuai dengan distribusi data. Temuan ini sejalan dengan karakteristik SVM yang sangat sensitif terhadap pemilihan parameter C dan gamma, terutama pada dataset berdimensi menengah dengan hubungan non-linear antar fitur.

Sebaliknya, *Random Forest* menunjukkan perubahan performa yang relatif terbatas setelah proses *tuning*. Hal ini mengindikasikan bahwa konfigurasi default *Random Forest* telah mampu menangkap struktur data secara cukup optimal, sehingga ruang peningkatan melalui optimasi parameter menjadi lebih sempit. Perbedaan respons terhadap *Hyperparameter tuning* ini menegaskan bahwa tidak semua algoritma memperoleh manfaat yang sama dari proses optimasi, dan efektivitas *tuning* sangat bergantung pada karakteristik model dan dataset yang digunakan.

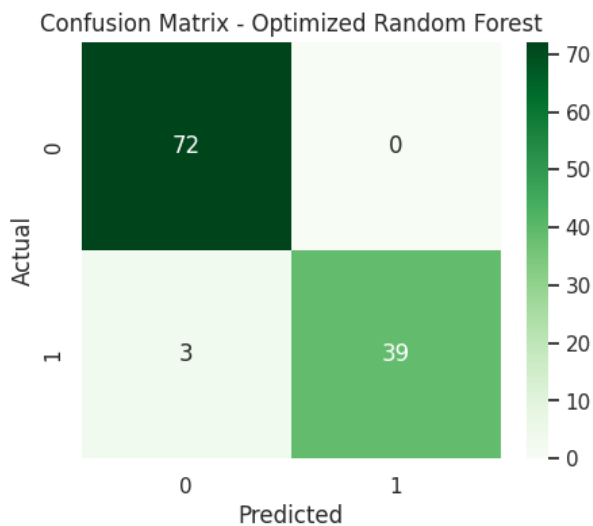
3.3 Analisis Distribusi Kesalahan Klasifikasi

Untuk memahami perilaku model secara lebih mendalam, dilakukan analisis distribusi kesalahan klasifikasi menggunakan *confusion matrix*. Visualisasi *confusion matrix* digunakan untuk menunjukkan jumlah prediksi benar dan salah pada masing-masing kelas, serta mengidentifikasi jenis kesalahan yang dihasilkan oleh model.



Gambar 2. *Confusion matrix* model SVM teroptimasi

Pada model SVM teroptimasi, *confusion matrix* menunjukkan bahwa sebagian besar sampel diklasifikasikan dengan benar. Jumlah *false negative* pada kelas malignant relatif rendah, yang menandakan bahwa model memiliki sensitivitas yang baik dalam mendeteksi kasus kanker payudara. Selain itu, tidak ditemukannya atau sangat minimnya *false positive* menunjukkan bahwa model juga mampu menghindari kesalahan diagnosis pada pasien non-kanker.



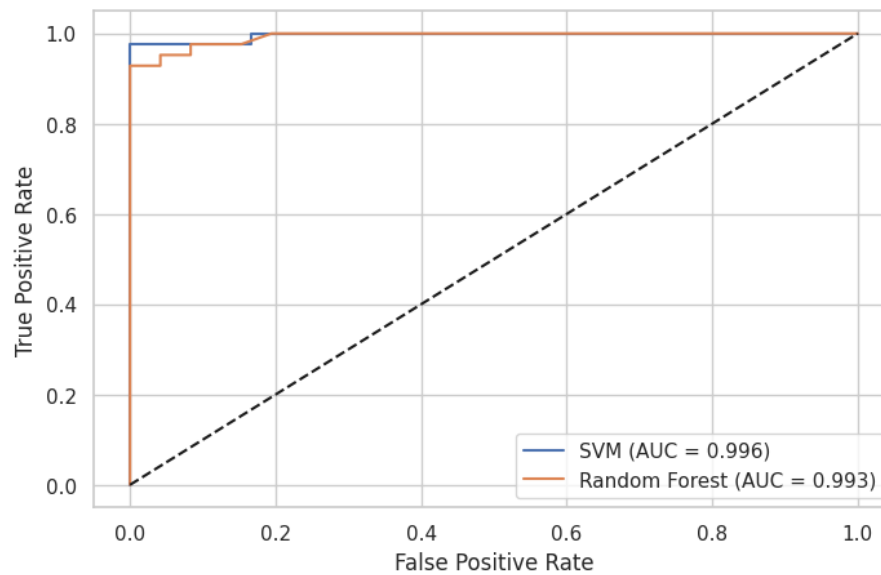
Gambar 3. *Confusion matrix* model *Random Forest* teroptimasi

Pada model *Random Forest* teroptimasi, *confusion matrix* menunjukkan pola yang serupa, namun dengan jumlah *false negative* yang sedikit lebih tinggi dibandingkan SVM. Meskipun perbedaannya tidak besar secara numerik, perbedaan ini menjadi signifikan dalam konteks medis, di mana setiap kasus kanker yang tidak terdeteksi dapat berdampak serius pada pasien. Oleh karena itu, analisis *confusion matrix* memberikan wawasan penting yang tidak dapat diperoleh hanya dari metrik akurasi.

Visualisasi *confusion matrix* untuk masing-masing model dapat disajikan sebagai Gambar 2 (SVM teroptimasi) dan Gambar 3 (*Random Forest* teroptimasi) untuk memperkuat analisis distribusi kesalahan klasifikasi.

3.4 Evaluasi Kemampuan Diskriminasi Model menggunakan ROC-AUC

Kemampuan diskriminasi model dianalisis lebih lanjut menggunakan kurva *Receiver Operating Characteristic* (ROC) dan nilai *Area Under the Curve* (AUC). Kurva ROC menggambarkan hubungan antara *true positive rate* dan *false positive rate* pada berbagai ambang keputusan, sehingga memberikan gambaran menyeluruh mengenai performa model di berbagai skenario klasifikasi.



Gambar 4. Kurva ROC perbandingan model SVM dan *Random Forest*

Hasil eksperimen menunjukkan bahwa baik SVM maupun *Random Forest* memiliki kurva ROC yang mendekati sudut kiri atas, yang menandakan kemampuan diskriminasi yang tinggi. Namun, model SVM teroptimasi menunjukkan nilai AUC yang sedikit lebih tinggi dibandingkan *Random Forest*. Perbedaan ini menunjukkan bahwa SVM memiliki kemampuan yang lebih stabil dalam membedakan antara kelas benign dan malignant pada berbagai threshold.

Kurva ROC perbandingan antara model SVM dan *Random Forest* teroptimasi dapat disajikan dalam Gambar 4. Visualisasi ini memberikan representasi yang intuitif mengenai keunggulan relatif masing-masing model dan mendukung temuan dari analisis *confusion matrix* serta *recall* kelas malignant.

3.5 Perbandingan Komprehensif Model *Baseline* dan Teroptimasi

Perbandingan menyeluruh antara model *baseline* dan model teroptimasi disajikan dalam Tabel 2, yang merangkum metrik performa utama untuk seluruh konfigurasi model. Dari tabel tersebut, terlihat bahwa *Hyperparameter tuning* memberikan peningkatan yang signifikan pada model SVM, terutama dalam hal *recall* kelas malignant dan AUC. Sementara itu, *Random Forest* menunjukkan performa yang relatif stabil sebelum dan sesudah *tuning*.

Temuan ini mengindikasikan bahwa algoritma berbasis margin seperti SVM lebih diuntungkan oleh proses optimasi *Hyperparameter* pada dataset tabular dengan margin kelas yang cukup jelas. Sebaliknya, *Random Forest* yang mengandalkan mekanisme ensemble cenderung mencapai performa optimal lebih cepat, sehingga tambahan optimasi memberikan dampak yang lebih terbatas.

3.6 Diskusi Metodologis

Dari perspektif metodologis, hasil penelitian ini menegaskan pentingnya evaluasi model yang tidak hanya berfokus pada akurasi, tetapi juga mempertimbangkan metrik lain yang relevan dengan konteks aplikasi. Dalam diagnosis kanker payudara, *recall* kelas malignant dan analisis *false negative* memiliki peran yang sangat penting, karena kesalahan jenis ini berpotensi menimbulkan risiko klinis yang serius.

Selain itu, penelitian ini menunjukkan bahwa *Hyperparameter tuning* merupakan langkah krusial dalam pengembangan model *machine learning*, khususnya untuk algoritma yang sensitif terhadap konfigurasi parameter. Tanpa optimasi, model dapat menghasilkan performa yang suboptimal meskipun akurasi awal terlihat tinggi. Oleh karena itu, pendekatan eksperimen yang sistematis dan terukur, seperti yang diterapkan dalam penelitian ini, sangat diperlukan untuk memperoleh model yang andal.

4. KESIMPULAN

Penelitian ini telah mengevaluasi dan membandingkan kinerja algoritma *Support Vector Machine* (SVM) dan *Random Forest* dalam klasifikasi kanker payudara menggunakan Dataset *Breast Cancer Wisconsin (Diagnostic)*, dengan fokus utama pada pengaruh *Hyperparameter tuning* terhadap performa dan stabilitas model. Berdasarkan hasil eksperimen, kedua algoritma menunjukkan kemampuan klasifikasi yang sangat baik, dengan tingkat akurasi sekitar 97% serta nilai *precision*, *recall*, dan F1-score yang relatif seimbang untuk kelas benign dan malignant. Hasil penelitian menunjukkan bahwa penerapan *Hyperparameter tuning* memberikan dampak yang berbeda pada masing-masing algoritma. Model SVM mengalami peningkatan performa yang lebih konsisten setelah proses optimasi, terutama pada metrik *recall* kelas malignant dan nilai ROC-AUC, yang mencerminkan kemampuan model dalam mendeteksi kasus kanker secara lebih sensitif dan stabil. Sebaliknya, *Random Forest* menunjukkan performa yang relatif tidak berubah

secara signifikan setelah optimasi, yang mengindikasikan bahwa konfigurasi parameter default telah mendekati kondisi optimal untuk karakteristik dataset yang digunakan. Analisis *confusion matrix* mengungkap bahwa kesalahan klasifikasi masih didominasi oleh *false negative* pada kelas malignant, meskipun jumlahnya relatif kecil. Temuan ini menegaskan bahwa evaluasi model dalam konteks diagnosis medis tidak dapat hanya bergantung pada akurasi, tetapi harus mempertimbangkan metrik yang berorientasi pada keselamatan pasien, khususnya sensitivitas terhadap deteksi kanker. Oleh karena itu, pemilihan metrik evaluasi yang tepat menjadi aspek penting dalam pengembangan sistem pendukung diagnosis berbasis *machine learning*. Secara keseluruhan, penelitian ini menegaskan bahwa *Hyperparameter tuning* memiliki peran penting dalam memastikan reliabilitas dan stabilitas model, meskipun tidak selalu menghasilkan peningkatan performa numerik yang signifikan pada *dataset benchmark*. Penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan beragam, mengintegrasikan fitur klinis atau data citra medis, serta menerapkan pendekatan explainable AI guna meningkatkan transparansi dan kepercayaan dalam penerapan model *machine learning* untuk diagnosis kanker payudara.

REFERENCES

- [1] I. Ahmad and F. Alqurashi, "Early cancer detection using deep learning and medical imaging: A survey," Dec. 01, 2024, *Elsevier Ireland Ltd.* doi: 10.1016/j.critrevonc.2024.104528.
- [2] S. Rezaei *et al.*, "Role of machine learning in molecular pathology for breast cancer: A review on gene expression profiling and RNA sequencing application," Sep. 01, 2025, *Elsevier Ireland Ltd.* doi: 10.1016/j.critrevonc.2025.104780.
- [3] J. Cox, A. Bhatti, and A. Atapour-Abarghouei, "Using artificial intelligence in the analysis of CT scans of the axillary nodes in breast cancer: A systematic review," *European Journal of Radiology Artificial Intelligence*, vol. 4, p. 100040, Dec. 2025, doi: 10.1016/j.ejrai.2025.100040.
- [4] Z. Mohamadi *et al.*, "Implementation of artificial intelligence in detection, classification, and prognostication of osteosarcoma utilizing different assessment techniques: a systematic review," Jan. 01, 2025, *Elsevier B.V.* doi: 10.1016/j.ibmed.2025.100250.
- [5] F. A. Altuhaifa, K. T. Win, and G. Su, "Predicting lung cancer survival based on clinical data using machine learning: A review," Oct. 01, 2023, *Elsevier Ltd.* doi: 10.1016/j.compbimed.2023.107338.
- [6] M. Sami and F. Sierra, "Using Machine Learning (ML) for Heat Transfer Coefficient (HTC) measurement in buildings: A systematic review," Aug. 01, 2025, *Elsevier Ltd.* doi: 10.1016/j.buildenv.2025.113220.
- [7] E. Pashaei, E. Pashaei, and N. Aydin, "Gene selection using hybrid binary black hole algorithm and modified binary particle swarm optimization," *Genomics*, vol. 111, no. 4, pp. 669–686, Jul. 2019, doi: 10.1016/j.ygeno.2018.04.004.
- [8] L. Caroprese, E. Vocaturo, and E. Zumpano, "Argumentation approaches for explainable AI in medical informatics," *Intelligent Systems with Applications*, vol. 16, Nov. 2022, doi: 10.1016/j.iswa.2022.200109.
- [9] E. Galli *et al.*, "Circulating blood biomarkers for minimal residual disease in hepatocellular carcinoma: A systematic review," Apr. 01, 2025, *W.B. Saunders Ltd.* doi: 10.1016/j.ctrv.2025.102908.
- [10] A. Procopio, G. Cesarelli, L. Donisi, A. Merola, F. Amato, and C. Cosentino, "Combined mechanistic modeling and machine-learning approaches in systems biology – A systematic literature review," Oct. 01, 2023, *Elsevier Ireland Ltd.* doi: 10.1016/j.cmpb.2023.107681.
- [11] R. Miranda, C. Alves, A. Abelha, and J. Machado, "Data Platforms for Real-time Insights in Healthcare: Systematic Review," in *Procedia Computer Science*, Elsevier B.V., 2023, pp. 826–831. doi: 10.1016/j.procs.2023.03.110.
- [12] N. Abbasi Dashtaki, M. CheshmehSohrabi, M. Pashootanzadeh, and H. Baradaran Kashani, "Retrieving and discovering new knowledge from documents' abstracts in scientific databases: Proposing a query-based abstractive summarization model," *International Journal of Information Management Data Insights*, vol. 5, no. 2, Dec. 2025, doi: 10.1016/j.jjime.2025.100366.
- [13] W. You, Z. Yang, and G. Ji, "PLS-based gene subset augmentation and tumor-specific gene identification," *Comput. Biol. Med.*, vol. 174, May 2024, doi: 10.1016/j.compbimed.2024.108434.

- [14] H. O. Al-Shamsi *et al.*, "Cancer research in the United Arab Emirates from birth to present: A bibliometric analysis," *Heliyon*, vol. 10, no. 6, Mar. 2024, doi: 10.1016/j.heliyon.2024.e27201.