

EVALUASI RECURSIVE FEATURE ELIMINATION UNTUK KLASIFIKASI KANKER PAYUDARA MENGGUNAKAN BERBAGAI ALGORITMA MACHINE LEARNING

Syarifah Yusnaini Putri¹, Sayuti rahman^{2,*}, Nia Ramadani³, Novalia Aprianti Ginting⁴,
Layla Syalsyadilla⁵, Dedi Agustriaman Zebua⁶

^{1,2,3,4}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Medan Area, Medan,
Sumatera Utara, Indonesia

^{5,6}Program Studi Magister Informatika, Pascasarjana, Universitas Medan Area, Medan,
Sumatera Utara, Indonesia

e-mail: ¹syarifahp25@gmail.com, ²sayutirahman@staff.uma.ac.id
³ramadaninia259@gmail.com, ⁴novaaliaaaa@gmail.com,
⁵laylasyalsyadilla@gmail.com, ⁶dedyagustriaman01@gmail.com

ABSTRAK

Deteksi dini kanker payudara membutuhkan model klasifikasi yang tidak hanya akurat, tetapi juga efisien dan mudah diinterpretasikan. Penelitian ini mengevaluasi pengaruh Recursive Feature Elimination (RFE) terhadap kinerja beberapa algoritma machine learning pada klasifikasi kanker payudara. Dataset yang digunakan adalah Wisconsin Diagnostic Breast Cancer (WDBC) dari UCI Machine Learning Repository yang terdiri atas 569 sampel dan 30 fitur numerik. Tahapan penelitian meliputi pra-pemrosesan data, penghapusan atribut noninformatif, standardisasi fitur menggunakan StandardScaler, pembagian data latih dan data uji dengan rasio 80:20, seleksi fitur menggunakan RFE berbasis Logistic Regression, serta pelatihan dan pengujian 11 algoritma klasifikasi. Kinerja model dievaluasi menggunakan akurasi, presisi, recall, F1-score, confusion matrix, dan kurva Receiver Operating Characteristic (ROC). Hasil pengujian menunjukkan bahwa sebelum seleksi fitur, Support Vector Machine, Logistic Regression, dan Voting Classifier memperoleh akurasi tertinggi sebesar 98,25%. Setelah penerapan RFE, akurasi ketiga model tersebut menjadi 97,37%, tetapi jumlah fitur berkurang dari 30 menjadi 15. Beberapa algoritma, seperti Nearest Centroid, Naïve Bayes, dan AdaBoost, mengalami peningkatan akurasi setelah RFE. Temuan ini menunjukkan bahwa RFE tidak selalu meningkatkan akurasi model terbaik, tetapi mampu menghasilkan model yang lebih ringkas, efisien, dan lebih mudah diinterpretasikan.

Kata kunci: kanker payudara; machine learning; RFE; klasifikasi; optimasi fitur.

ABSTRACT

Early detection of breast cancer requires classification models that are not only accurate but also efficient and interpretable. This study evaluates the effect of Recursive Feature Elimination (RFE) on the performance of several machine learning algorithms for breast cancer classification. The dataset used is the Wisconsin Diagnostic Breast Cancer (WDBC) dataset from the UCI Machine Learning Repository, consisting of 569 samples and 30 numerical features. The research stages include data preprocessing, removal of non-informative attributes, feature standardization using StandardScaler, train-test splitting with an 80:20 ratio, feature selection using Logistic Regression-based RFE, and training and testing of 11 classification algorithms. Model performance was evaluated using accuracy, precision, recall, F1-score, confusion matrix, and

Receiver Operating Characteristic (ROC) curve. The results show that before feature selection, Support Vector Machine, Logistic Regression, and Voting Classifier achieved the highest accuracy of 98.25%. After applying RFE, the accuracy of these models decreased slightly to 97.37%, while the number of features was reduced from 30 to 15. Several algorithms, including Nearest Centroid, Naïve Bayes, and AdaBoost, showed improved accuracy after RFE. These findings indicate that RFE does not always improve the best model accuracy, but it can produce a more compact, efficient, and interpretable classification model.

Keywords: *breast cancer; machine learning; recursive feature elimination (RFE); classification; feature optimization*

1. PENDAHULUAN

Kanker payudara masih menjadi salah satu masalah kesehatan global karena jumlah kasus baru dan dampak mortalitasnya yang tinggi. Laporan GLOBOCAN menunjukkan bahwa kanker payudara merupakan salah satu jenis kanker dengan jumlah kasus tertinggi di dunia, sehingga deteksi dini menjadi bagian penting dalam upaya peningkatan peluang kesembuhan pasien [1], [2]. Secara klinis, kanker payudara dapat berkembang dari pertumbuhan sel abnormal pada jaringan payudara dan dapat bersifat ganas apabila mampu menyebar ke jaringan atau organ lain [3].

Diagnosis kanker payudara umumnya dilakukan melalui pemeriksaan klinis, pencitraan medis, dan pemeriksaan patologi seperti biopsi. Meskipun prosedur tersebut merupakan standar dalam praktik medis, proses diagnosis tetap memerlukan waktu, sumber daya, dan keahlian tenaga medis yang memadai [4]. Oleh karena itu, sistem pendukung keputusan berbasis komputasi semakin banyak dikembangkan untuk membantu proses klasifikasi dan penyaringan awal, terutama pada data numerik atau citra medis yang memiliki pola kompleks [5], [6].

Perkembangan machine learning memberikan peluang untuk membangun model klasifikasi yang mampu mengenali pola data kanker payudara secara otomatis. Sejumlah penelitian menunjukkan bahwa algoritma seperti Support Vector Machine (SVM), Logistic Regression, Random Forest, K-Nearest Neighbor (KNN), Decision Tree, AdaBoost, dan model ensemble dapat memberikan performa yang baik pada data diagnosis kanker payudara [7]-[10]. Namun, tingginya jumlah fitur dan adanya korelasi antar fitur dapat meningkatkan kompleksitas model, memperbesar risiko redundansi informasi, serta mengurangi interpretabilitas hasil.

Salah satu dataset yang sering digunakan dalam penelitian klasifikasi kanker payudara adalah Wisconsin Diagnostic Breast Cancer (WDBC). Dataset ini berisi karakteristik numerik hasil ekstraksi citra fine needle aspirate (FNA) dan telah banyak digunakan sebagai benchmark dalam evaluasi model klasifikasi [11]. Pada dataset semacam ini, seleksi fitur menjadi penting karena tidak semua fitur memberikan kontribusi yang sama terhadap proses prediksi. Penggunaan seluruh fitur memang dapat menghasilkan akurasi tinggi, tetapi model yang terlalu kompleks belum tentu lebih efisien atau lebih mudah dijelaskan.

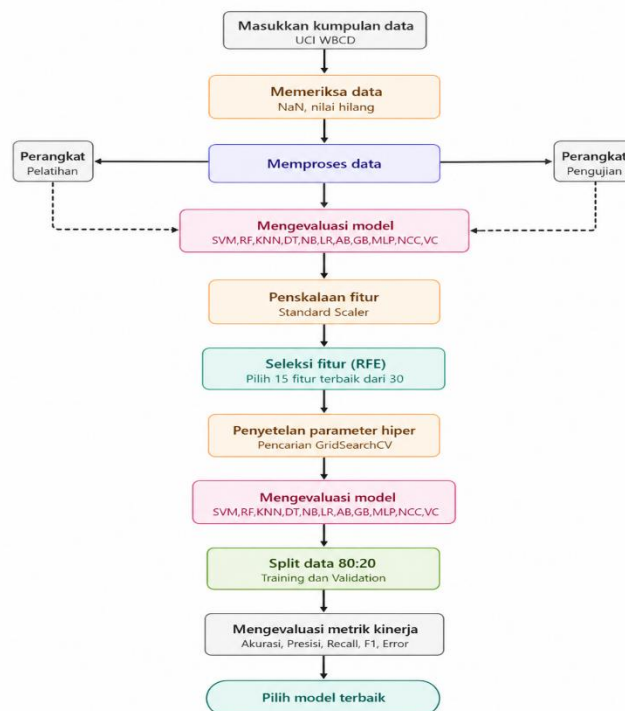
Recursive Feature Elimination (RFE) merupakan metode seleksi fitur yang bekerja secara iteratif dengan melatih estimator, menghitung bobot kepentingan fitur, menghapus fitur dengan kontribusi terendah, dan mengulangi proses tersebut hingga jumlah fitur yang diinginkan tercapai [12], [13]. Berbeda dari reduksi dimensi seperti Principal Component Analysis (PCA), RFE tetap mempertahankan nama fitur asli sehingga hasilnya lebih mudah diinterpretasikan. Dalam konteks diagnosis kanker payudara, aspek

interpretabilitas ini penting karena model yang ringkas dapat membantu menjelaskan fitur apa saja yang paling berperan dalam klasifikasi [14],[15].

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengevaluasi pengaruh RFE terhadap kinerja 11 algoritma machine learning pada klasifikasi kanker payudara menggunakan dataset WDBC. Kontribusi utama penelitian ini adalah membandingkan performa model sebelum dan sesudah seleksi fitur, mengidentifikasi algoritma yang paling stabil setelah pengurangan fitur, serta menilai trade-off antara akurasi dan efisiensi model. Dengan demikian, penelitian ini tidak hanya berfokus pada pencapaian akurasi tertinggi, tetapi juga pada penyederhanaan model agar lebih efisien dan mudah diinterpretasikan.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen komputasional untuk membandingkan performa beberapa algoritma machine learning sebelum dan sesudah penerapan seleksi fitur RFE. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur metodologi penelitian

Dataset Penelitian

Dataset yang digunakan adalah Wisconsin Diagnostic Breast Cancer (WDBC) yang tersedia pada UCI Machine Learning Repository. Dataset ini terdiri atas 569 sampel dengan 30 fitur numerik dan satu label kelas diagnosis, yaitu jinak (benign) dan ganas

(malignant) [7]. Fitur pada dataset merepresentasikan karakteristik inti sel, seperti radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, dan fractal dimension. Setiap karakteristik dihitung dalam tiga bentuk, yaitu nilai rata-rata (mean), galat standar (standard error), dan nilai terburuk (worst). Atribut id tidak digunakan dalam pelatihan model karena tidak memiliki makna prediktif terhadap kelas diagnosis.

Pra-pemrosesan Data

Pra-pemrosesan dilakukan melalui beberapa tahap. Pertama, atribut noninformatif dihapus dari dataset. Kedua, label diagnosis dikodekan ke dalam bentuk numerik agar dapat diproses oleh algoritma machine learning. Ketiga, fitur numerik distandardisasi menggunakan StandardScaler sehingga setiap fitur memiliki rata-rata mendekati nol dan simpangan baku satu. Standardisasi diperlukan karena beberapa algoritma, seperti SVM, KNN, Logistic Regression, dan MLP, sensitif terhadap skala fitur. Data kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20 menggunakan stratified sampling agar distribusi kelas pada data latih dan data uji tetap seimbang.

Recursive Feature Elimination

RFE digunakan untuk memilih 15 fitur paling relevan dari 30 fitur awal. Dalam penelitian ini, Logistic Regression digunakan sebagai estimator karena memiliki koefisien model yang dapat digunakan untuk menentukan tingkat kepentingan fitur. Proses RFE dilakukan secara bertahap dengan melatih estimator, menghitung bobot fitur, menghapus fitur dengan bobot terendah, kemudian mengulang proses tersebut sampai tersisa 15 fitur. Secara sederhana, bobot kepentingan fitur ke- j dapat dinyatakan pada Persamaan (1).

$$S_j = |\beta_j|, j = 1, 2, \dots, p \tag{1}$$

dengan S_j menyatakan skor kepentingan fitur ke- j , β_j merupakan koefisien model, dan p adalah jumlah fitur awal. Setelah semua fitur diberi skor, fitur diurutkan berdasarkan skor tertinggi. Subset fitur terpilih dinyatakan pada Persamaan (2).

$$F_k = f_{(1)}, f_{(2)}, \dots, f_{(k)}, \text{ dengan } S_{(1)} \geq S_{(2)} \geq \dots \geq S_{(p)} \tag{2}$$

Pada penelitian ini, nilai k ditetapkan sebesar 15 sehingga model akhir menggunakan 50% fitur dari dataset awal. Pemilihan 15 fitur bertujuan menilai apakah model tetap mampu mempertahankan performa yang kompetitif setelah jumlah fitur dikurangi.

Tabel 1. Fitur yang dipilih oleh RFE

No.	Fitur Terpilih	Keterangan
1	area_mean	Luas rata-rata inti sel
2	compactness_mean	Kekompakan rata-rata bentuk inti sel
3	concave_points_mean	Jumlah titik cekung rata-rata
4	radius_se	Galat standar radius
5	perimeter_se	Galat standar perimeter
6	area_se	Galat standar luas
7	compactness_se	Galat standar kekompakan

8	radius_worst	Nilai radius terburuk
9	texture_worst	Nilai tekstur terburuk
10	perimeter_worst	Nilai perimeter terburuk
11	area_worst	Nilai luas terburuk
12	smoothness_worst	Nilai kehalusan terburuk
13	concavity_worst	Nilai cekungan terburuk
14	concave_points_worst	Nilai titik cekung terburuk
15	symmetry_worst	Nilai simetri terburuk

Algoritma Klasifikasi dan Evaluasi

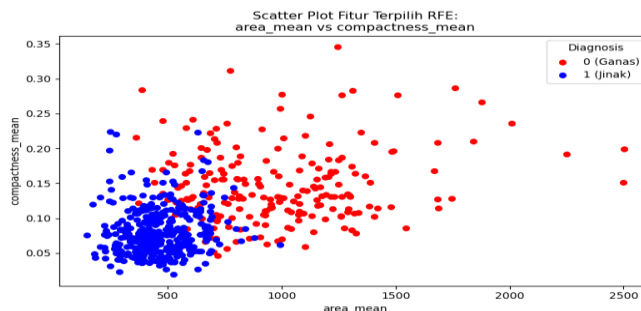
Sebanyak 11 algoritma klasifikasi digunakan dalam penelitian ini, yaitu Random Forest, Support Vector Machine, Nearest Centroid, Naïve Bayes, Multilayer Perceptron, Logistic Regression, Decision Tree, AdaBoost, K-Nearest Neighbor, Gradient Boosting, dan Voting Classifier. Evaluasi dilakukan pada dua skenario. Skenario pertama menggunakan seluruh 30 fitur sebagai baseline, sedangkan skenario kedua menggunakan 15 fitur seleksi RFE. Agar perbandingan adil, pembagian data dan prosedur standardisasi dijaga tetap sama pada kedua skenario.

Kinerja model diukur menggunakan akurasi, presisi, recall, dan F1-score. Akurasi menunjukkan proporsi prediksi yang benar terhadap seluruh data uji. Presisi mengukur ketepatan prediksi pada kelas positif, recall mengukur kemampuan model menemukan kelas positif, sedangkan F1-score merupakan rata-rata harmonik antara presisi dan recall. Selain itu, confusion matrix digunakan untuk mengamati jenis kesalahan klasifikasi, dan kurva ROC digunakan untuk melihat kemampuan model membedakan kelas.

3. HASIL DAN PEMBAHASAN

Visualisasi Awal Data

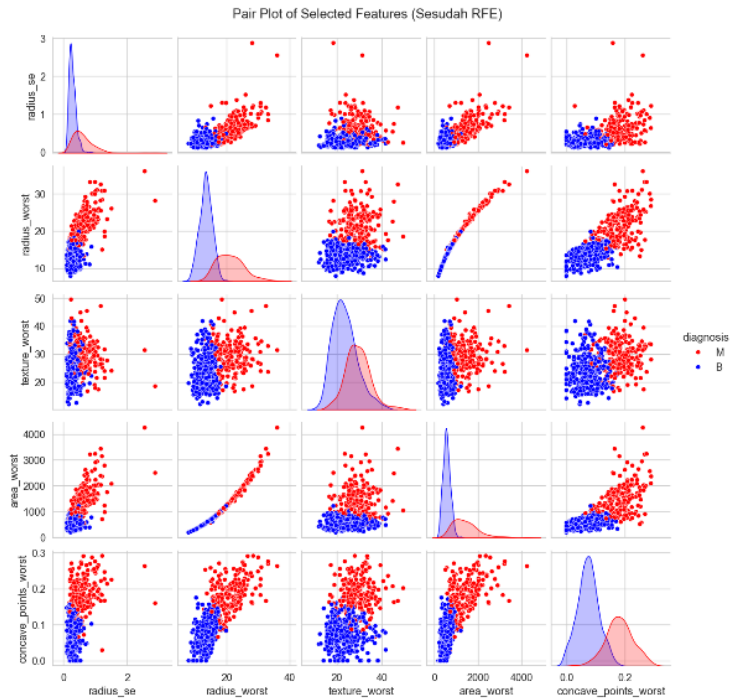
Visualisasi awal dilakukan untuk memahami karakteristik data dan indikasi pemisahan kelas. Gambar 2 memperlihatkan hubungan antara *area_mean* dan *compactness_mean* sebagai contoh fitur yang terpilih oleh RFE. Secara umum, sampel ganas cenderung memiliki area dan kekompakan yang lebih tinggi dibandingkan sampel jinak, meskipun masih terdapat area tumpang tindih antarkelas.



Gambar 2. Diagram sebar dari fitur (*smoothness_mean*, *area_mean*)

Gambar 3 memperlihatkan hubungan pasangan beberapa fitur utama. Pola distribusi menunjukkan adanya korelasi antar fitur, khususnya pada *radius_mean*, *perimeter_mean*,

dan area_mean. Korelasi yang tinggi antar fitur menjadi salah satu alasan perlunya seleksi fitur agar model tidak menggunakan informasi yang terlalu redundan.



Gambar 3. Diagram pasangan dari Fitur (radius_mean, texture_mean, perimeter_mean, area_mean, smoothness_mean)

Hasil Klasifikasi Sebelum Seleksi Fitur

Pengujian awal dilakukan menggunakan seluruh 30 fitur. Hasil pada Tabel 2 menunjukkan bahwa SVM, Logistic Regression, dan Voting Classifier memperoleh akurasi tertinggi sebesar 98,25%, dengan presisi, recall, dan F1-score sebesar 97,62%. Hasil ini menunjukkan bahwa fitur awal pada dataset WDBC sudah memiliki kemampuan diskriminatif yang kuat untuk membedakan kelas jinak dan ganas.

Tabel 2. Pengujian sebelum seleksi fitur

Metode	Presisi	Recall	F-Measure	Akurasi	Error
Random Forest	95.12	92.86	93.98	95.61	4.386
Support Vector Machine	97.62	97.62	97.62	98.25	1.754
Nearest Centroid	94.87	88.1	91.36	93.86	6.14
Naïve Bayes	90.48	90.48	90.48	92.98	7.018
Multilayer Perceptron	93.18	97.62	95.35	96.49	3.509
Logistic Regression	97.62	97.62	97.62	98.25	1.754
Decision Tree	84.78	92.86	88.64	91.23	8.772
AdaBoost	97.44	90.48	93.83	95.61	4.386
K-Nearest Neighbor	95.12	92.86	93.83	95.61	4.386
Gradient Boosting	97.44	90.48	93.83	95.61	4.386

Voting Classifier	97.62	97.62	97.62	98.25	1.754
-------------------	-------	-------	-------	-------	-------

Hasil Klasifikasi Setelah Seleksi Fitur RFE

Setelah jumlah fitur direduksi dari 30 menjadi 15 menggunakan RFE, dilakukan pelatihan ulang pada seluruh algoritma. Hasil pada Tabel 3 menunjukkan bahwa SVM, Logistic Regression, AdaBoost, dan Voting Classifier memperoleh akurasi sebesar 97,37%. Meskipun nilai ini sedikit lebih rendah dibandingkan akurasi tertinggi sebelum RFE, model tetap mempertahankan performa yang tinggi dengan jumlah fitur yang lebih sedikit.

Tabel 3. Kinerja model setelah seleksi fitur RFE

Metode	Presisi	Recall	F-Measure	Akurasi	Error
Random Forest	95.12	92.86	93.98	95.61	4.386
Support Vector Machine	95.35	97.62	96.47	97.37	2.632
Nearest Centroid	100	90.48	95	96.49	3.509
Naïve Bayes	95.12	92.86	93.98	95.61	4.386
Multilayer Perceptron	91.11	97.62	94.25	95.61	4.386
Logistic Regression	95.35	97.62	96.47	97.37	2.632
Decision Tree	83.33	95.24	88.89	91.23	8.772
AdaBoost	100	92.86	96.3	97.37	2.632
K-Nearest Neighbor	95.12	92.86	93.98	95.61	4.386
Gradient Boosting	95.12	92.86	93.98	95.61	4.386
Voting Classifier	95.35	97.62	96.47	97.37	2.632

Perbandingan Sebelum dan Sesudah RFE

Perbandingan akurasi pada Tabel 4 memperlihatkan bahwa RFE memberikan dampak yang berbeda pada setiap algoritma. Nearest Centroid dan Naïve Bayes mengalami peningkatan akurasi sebesar 2,63%, sedangkan AdaBoost meningkat sebesar 1,76%. Sementara itu, SVM, Logistic Regression, Multilayer Perceptron, dan Voting Classifier mengalami penurunan sebesar 0,88%. Penurunan kecil ini menunjukkan bahwa sebagian fitur yang dieliminasi masih membawa informasi prediktif bagi model yang sejak awal sudah kuat.

Tabel 4. Kinerja model setelah seleksi fitur RFE

Methods	Sebelum RFE (%)	Sesudah RFE (%)	Selisih (%)
Random Forest	95.61	95.61	0
Support Vector Machine	98.25	97.37	-0.88
Nearest Centroid	93.86	96.49	2.63
Naïve Bayes	92.98	95.61	2.63
Multilayer Perceptron	96.49	95.61	-0.88

Logistic Regression	98.25	97.37	-0.88
Decision Tree	91.23	91.23	0
AdaBoost	95.61	97.37	1.76
K-Nearest Neighbor	95.61	95.61	0
Gradient Boosting	95.61	95.61	0
Voting Classifier	98.25	97.37	-0.88

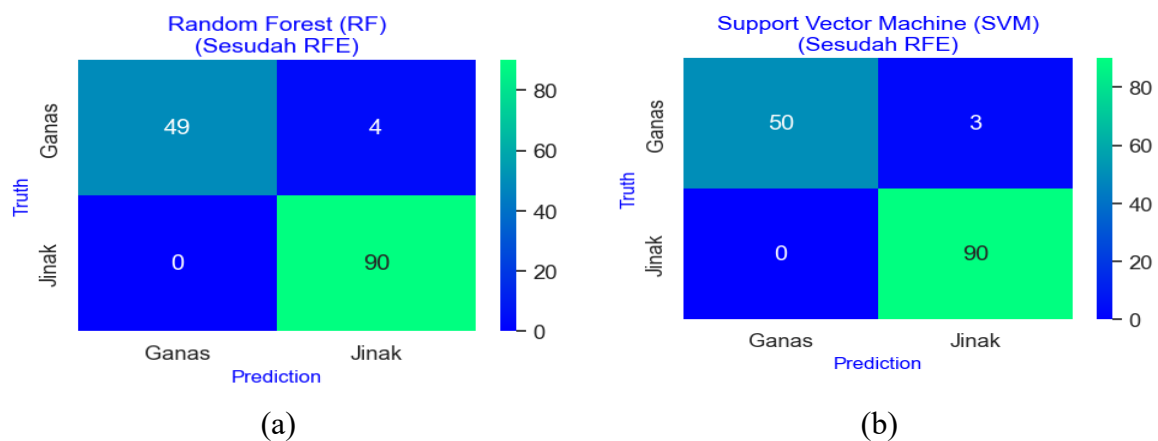
Dengan menggunakan 15 fitur, model setelah RFE tetap mampu mencapai akurasi maksimum 97,37%. Artinya, pengurangan fitur sebesar 50% hanya menurunkan akurasi terbaik sebesar 0,88%. Dalam konteks implementasi sistem pendukung keputusan, kondisi ini dapat diterima apabila tujuan utama tidak hanya mengejar akurasi tertinggi, tetapi juga memperoleh model yang lebih ringkas dan mudah dijelaskan.

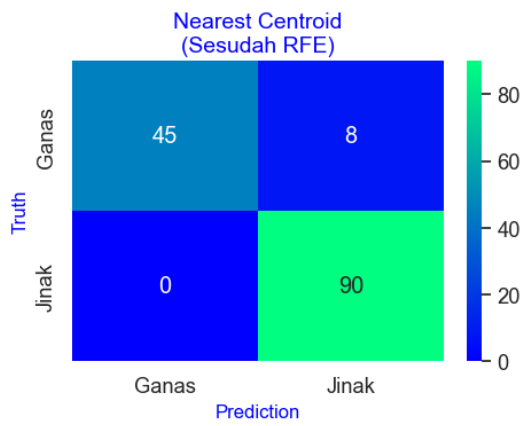
Pada algoritma Nearest Centroid dan Naïve Bayes, RFE justru meningkatkan akurasi. Hal ini mengindikasikan bahwa kedua model tersebut lebih sensitif terhadap fitur yang redundan atau kurang relevan. Dengan mengurangi fitur, batas keputusan menjadi lebih sederhana dan performa model meningkat. Sebaliknya, model seperti SVM dan Logistic Regression sudah mampu memanfaatkan informasi dari seluruh fitur dengan baik, sehingga penghapusan sebagian fitur dapat mengurangi informasi yang diperlukan untuk klasifikasi.

Voting Classifier memiliki performa terbaik pada skenario baseline dan tetap kompetitif setelah RFE. Hal ini menunjukkan bahwa pendekatan ensemble mampu menjaga stabilitas prediksi karena menggabungkan beberapa model dasar. Namun, karena akurasi tertinggi tetap diperoleh sebelum RFE, hasil penelitian ini menegaskan bahwa seleksi fitur perlu dievaluasi berdasarkan tujuan analisis. Jika target utama adalah akurasi maksimal, penggunaan seluruh fitur masih lebih unggul. Namun, jika targetnya adalah model yang lebih sederhana dan efisien, RFE menjadi pilihan yang layak.

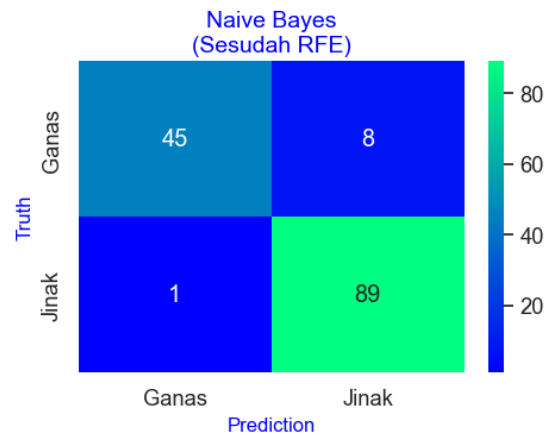
Analisis Confusion Matrix dan ROC

Confusion matrix pada Gambar 4 menunjukkan bahwa sebagian besar model mampu mengklasifikasikan sampel jinak dan ganas dengan jumlah kesalahan yang relatif rendah. SVM dan Voting Classifier setelah RFE menghasilkan tiga kesalahan pada kelas ganas dan tidak menghasilkan kesalahan pada kelas jinak. Decision Tree memiliki jumlah kesalahan paling tinggi, sehingga performanya relatif lebih rendah dibandingkan model lain.

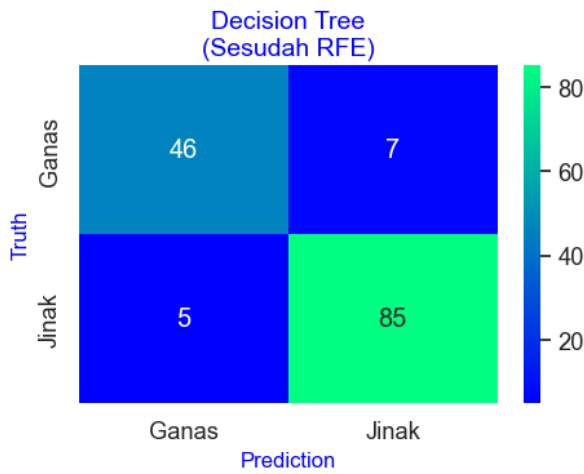




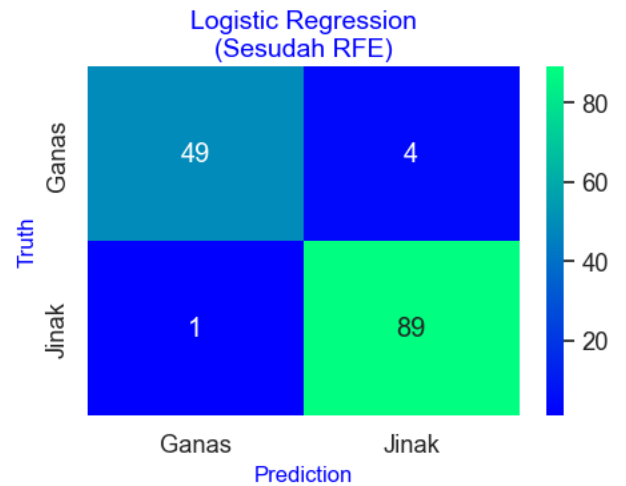
(c)



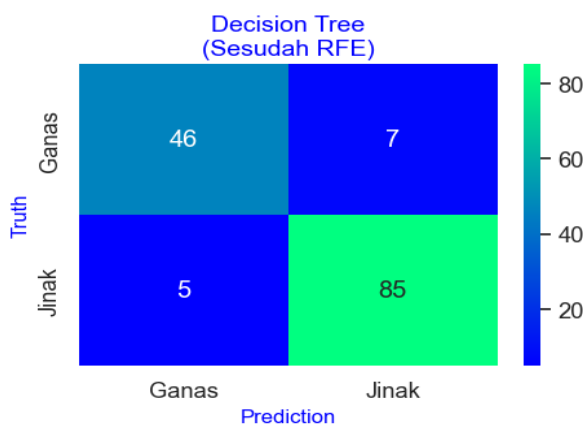
(d)



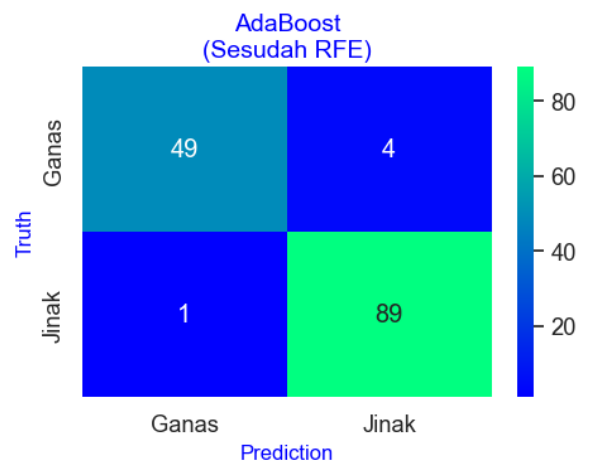
(e)



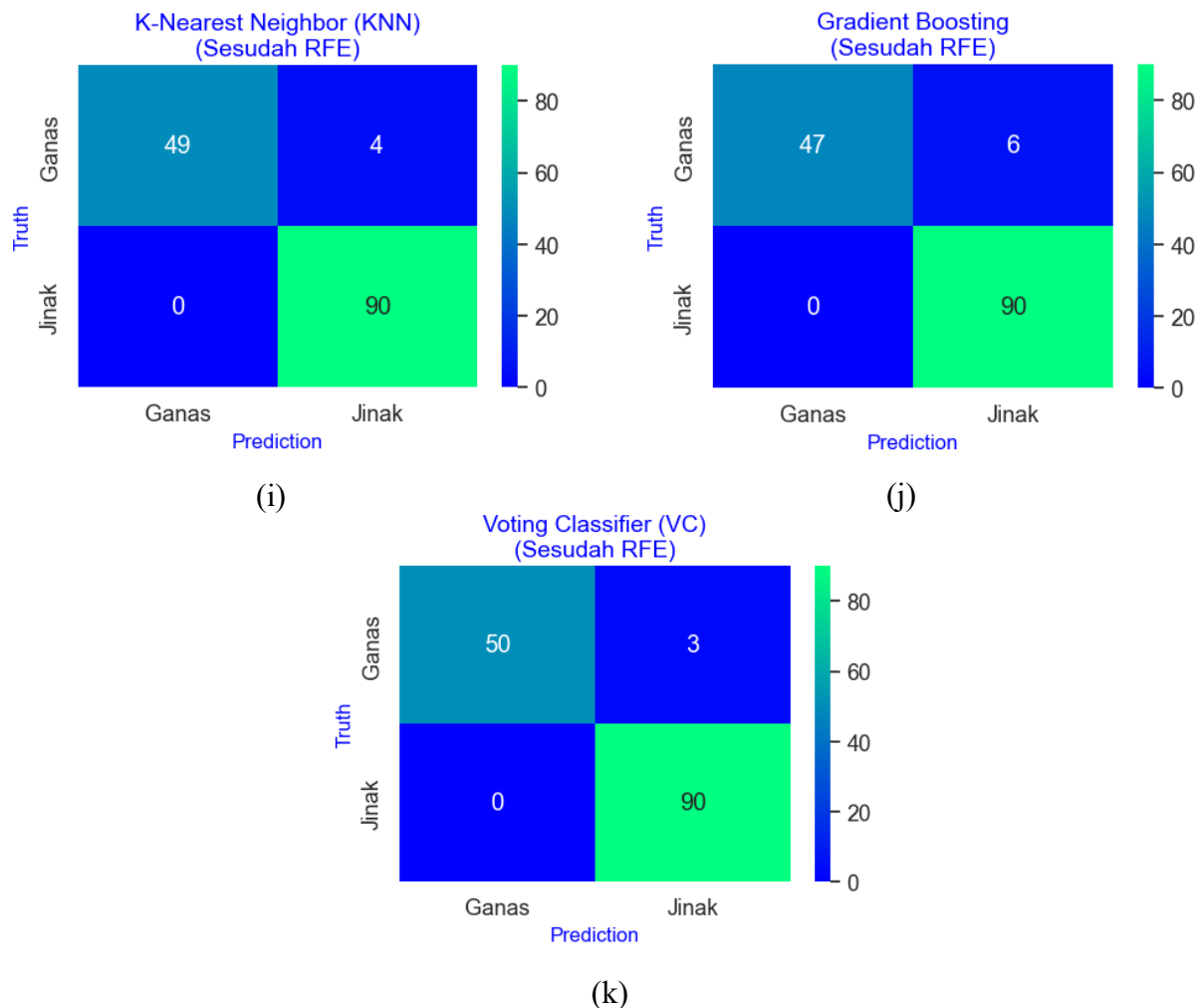
(f)



(g)



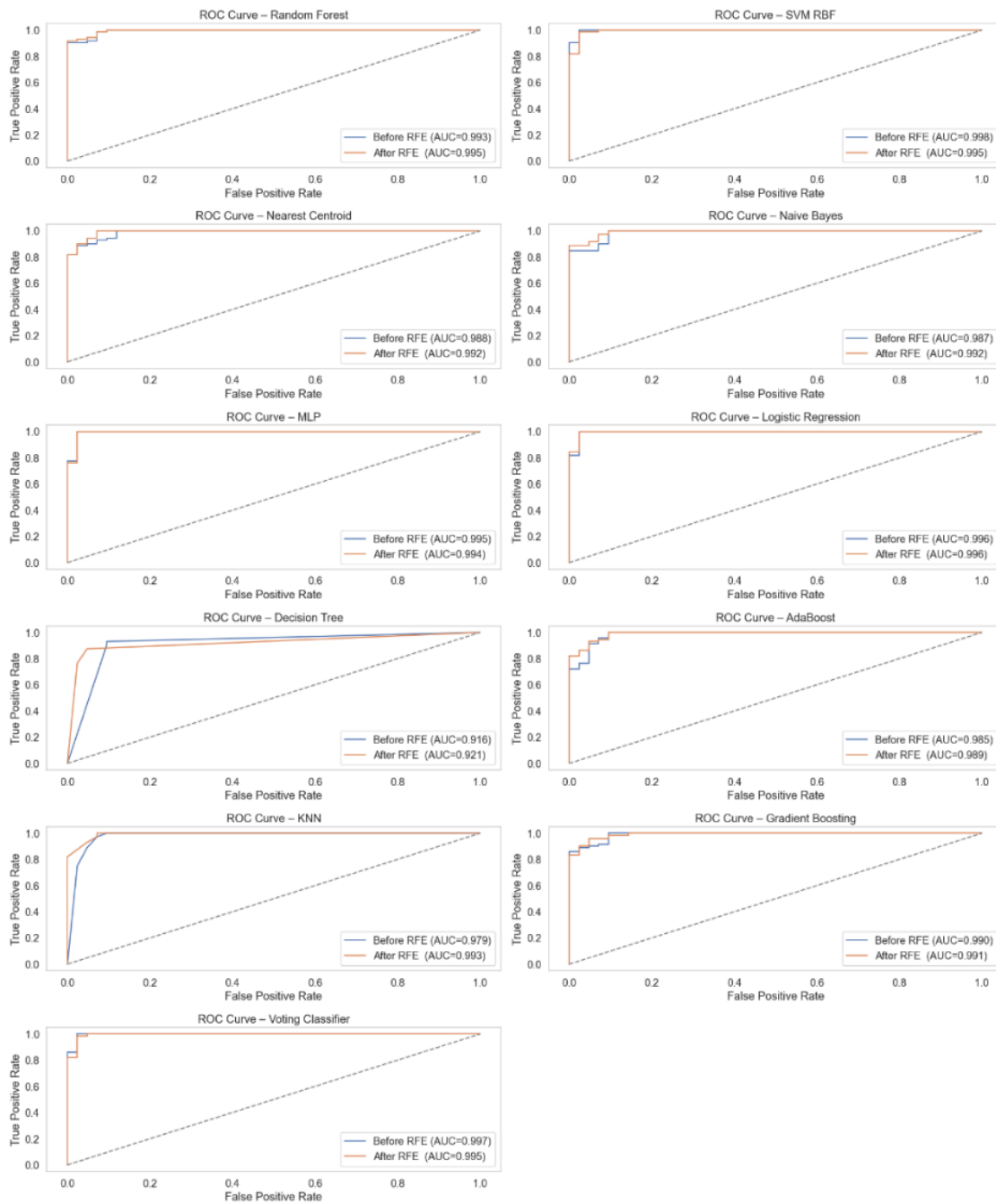
(h)



Gambar 4. Matriks kebingungan untuk (a) Random Forest (b) Support Vector Machine (c) Nearest Centroid (d) Naïve Bayes (e) Multilayer perceptron (f) Logistic Regression (g) Decision Tree (h) AdaBoost (i) K-Nearest Neighbor (j) Gradient Boosting (k) Voting Classifier

Kurva ROC pada Gambar 5 menunjukkan bahwa sebagian besar model memiliki nilai AUC yang tinggi. Hal ini mengindikasikan kemampuan model dalam membedakan kelas jinak dan ganas secara baik. Namun, evaluasi ROC perlu dibaca bersama confusion matrix dan metrik klasifikasi lain agar interpretasi performa tidak hanya bergantung pada satu indikator.

ROC Curve Before vs After RFE Optimization



Gambar 7. Kurva ROC model klasifikasi sebelum dan sesudah optimasi fitur.

Implikasi dan Keterbatasan

Hasil penelitian menunjukkan bahwa RFE bermanfaat ketika peneliti membutuhkan model yang lebih sederhana tanpa kehilangan performa secara signifikan. Fitur yang lebih

sedikit dapat mempercepat proses pelatihan, mengurangi kompleksitas komputasi, dan memudahkan interpretasi fitur yang berkontribusi terhadap klasifikasi.

Penelitian ini memiliki beberapa keterbatasan. Pertama, evaluasi dilakukan pada satu dataset benchmark sehingga generalisasi ke data klinis nyata masih perlu diuji. Kedua, jumlah fitur yang dipilih ditetapkan sebanyak 15, sehingga penelitian lanjutan dapat menggunakan RFECV untuk menentukan jumlah fitur optimal secara otomatis. Ketiga, penelitian ini belum membahas validasi eksternal dan optimasi hiperparameter secara mendalam. Oleh karena itu, pengembangan berikutnya dapat menambahkan validasi pada dataset lain, pencarian hiperparameter yang lebih sistematis, dan analisis interpretabilitas menggunakan metode seperti SHAP atau permutation importance.

4. KESIMPULAN

Penelitian ini mengevaluasi pengaruh Recursive Feature Elimination terhadap kinerja 11 algoritma machine learning dalam klasifikasi kanker payudara menggunakan dataset WDBC. Hasil pengujian menunjukkan bahwa sebelum seleksi fitur, SVM, Logistic Regression, dan Voting Classifier mencapai akurasi tertinggi sebesar 98,25%. Setelah penerapan RFE, akurasi tertinggi menjadi 97,37%, tetapi jumlah fitur berhasil dikurangi dari 30 menjadi 15. Temuan ini menunjukkan bahwa RFE tidak selalu meningkatkan akurasi model terbaik, terutama pada algoritma yang sejak awal sudah memiliki performa tinggi. Namun, RFE mampu meningkatkan performa beberapa algoritma seperti Nearest Centroid, Naïve Bayes, dan AdaBoost, serta menghasilkan model yang lebih ringkas dan mudah diinterpretasikan. Dengan demikian, RFE layak digunakan sebagai teknik seleksi fitur ketika tujuan penelitian tidak hanya mengejar akurasi tertinggi, tetapi juga mempertimbangkan efisiensi dan interpretabilitas model.

Penelitian lanjutan disarankan untuk menggunakan validasi eksternal, menentukan jumlah fitur optimal menggunakan RFECV, melakukan optimasi hiperparameter yang lebih sistematis, serta menguji model pada dataset klinis yang lebih besar dan lebih beragam.

DAFTAR PUSTAKA

- [1] H. Sung et al., “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [2] L. Wilkinson and T. Gathani, “Understanding breast cancer as a global health concern,” *British Journal of Radiology*, vol. 95, no. 1130, p. 20211033, 2022.
- [3] A. Stanislawek, “Breast cancer—epidemiology, risk factors, classification, prognostic markers, and current treatment strategies: An updated review,” *Cancers*, vol. 13, no. 17, pp. 1–30, 2021.
- [4] M. Xiao et al., “Diagnostic value of breast lesions between deep learning-based computer-aided diagnosis system and experienced radiologists,” *Frontiers in Oncology*, vol. 10, pp. 1–10, 2020.
- [5] X. Y. Liew, N. Hameed, and J. Clos, “A review of computer-aided expert systems for breast cancer diagnosis,” *Cancers*, vol. 13, no. 11, p. 2764, 2021.
- [6] K. Loizidou, R. Elia, and C. Pitris, “Computer-aided breast cancer detection and

-
- classification in mammography: A comprehensive review,” *Computers in Biology and Medicine*, vol. 153, p. 106554, 2023.
- [7] A. Rasool et al., “Improved machine learning-based predictive models for breast cancer diagnosis,” *International Journal of Environmental Research and Public Health*, vol. 19, no. 6, p. 3211, 2022.
- [8] C. G. Yedjou et al., “Application of machine learning algorithms in breast cancer diagnosis and classification,” *International Journal of Science and Academic Research*, vol. 2, no. 1, pp. 3081–3086, 2021.
- [9] K. M. M. Uddin et al., “Machine learning-based diagnosis of breast cancer utilizing feature optimization technique,” *Computer Methods and Programs in Biomedicine Update*, vol. 3, p. 100098, 2023.
- [10] L. Alkhatlan and A. K. J. Saudagar, “Predicting and classifying breast cancer using machine learning,” *Journal of Computational Biology*, vol. 29, no. 6, pp. 497–514, 2022.
- [11] W. H. Wolberg, W. N. Street, and O. L. Mangasarian, “Breast Cancer Wisconsin (Diagnostic),” *UCI Machine Learning Repository*, 1993.
- [12] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, “Gene selection for cancer classification using support vector machines,” *Machine Learning*, vol. 46, pp. 389–422, 2002.
- [13] R. J. Urbanowicz et al., “Relief-based feature selection: Introduction and review,” *Journal of Biomedical Informatics*, vol. 85, pp. 189–203, 2018.
- [14] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [15] T. A. Assegie, R. L. Tulasi, and N. K. Kumar, “Breast cancer prediction model with decision tree and adaptive boosting,” *International Journal of Artificial Intelligence*, vol. 10, no. 1, pp. 184–190, 2021.