

Analisis Komparatif Metode Ensemble Learning pada Prediksi Kanker Payudara

Nur Cahyo IA^{1*}, Setyawan Wibisono²

¹².Program Studi Teknik Informatika,Fakultas Teknik dan Informatika, Universitas Stikubank
Semarang

e-mail: [*nurcahyoigbalardimas@mhs.unisbank.ac.id](mailto:nurcahyoigbalardimas@mhs.unisbank.ac.id)¹, setyawan@edu.unisbank.ac.id

Abstrak

Penelitian ini bertujuan untuk menganalisis dan membandingkan performa tiga metode ensemble learning, yaitu Random Forest, AdaBoost, dan XGBoost, dalam mengklasifikasikan data kanker payudara. Dataset yang digunakan adalah Wisconsin Breast Cancer (WBC) dari UCI Machine Learning Repository. Metode ensemble diterapkan karena kemampuannya dalam meningkatkan stabilitas dan akurasi prediksi dengan menggabungkan beberapa model dasar. Evaluasi dilakukan dengan menggunakan metrik akurasi, presisi, recall, F1-score, serta confusion matrix dan AUC, dengan pendekatan validasi silang 10-fold. Hasil penelitian menunjukkan bahwa AdaBoost memberikan hasil paling seimbang dengan akurasi tertinggi sebesar 96,31% dan kesalahan klasifikasi paling minimal. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem deteksi dini kanker berbasis machine learning.

Kata kunci— Ensemble Learning, Breast Cancer, AdaBoost, Random Forest, XGBoost

Abstract

This study aims to analyze and compare the performance of three ensemble learning methods, namely Random Forest, AdaBoost, and XGBoost, in classifying breast cancer data. The dataset used is Wisconsin Breast Cancer (WBC) from the UCI Machine Learning Repository. Ensemble methods were applied due to their ability to enhance prediction stability and accuracy by combining several base models. The models were evaluated using accuracy, precision, recall, F1-score, confusion matrix, and AUC metrics with a 10-fold cross-validation approach. The results show that AdaBoost performed the best with the highest accuracy of 96.31% and the lowest misclassification. This research is expected to contribute to the development of early detection systems for breast cancer using machine learning.

Keywords— Ensemble Learning, Breast Cancer, AdaBoost, Random Forest, XGBoost

1. PENDAHULUAN

Kanker payudara merupakan salah satu jenis kanker dengan tingkat prevalensi dan kematian yang tinggi di Indonesia. Berdasarkan data dari World Health Organization (WHO, 2024), kanker payudara menempati peringkat pertama sebagai jenis kanker terbanyak di Indonesia, dengan estimasi kasus baru mencapai lebih dari 65.000 setiap tahunnya. Tantangan utama dalam penanganan kanker payudara adalah keterlambatan deteksi yang menyebabkan pasien datang dalam kondisi stadium lanjut. Oleh karena itu, diperlukan strategi deteksi dini yang efektif guna menekan angka morbiditas dan mortalitas.

Seiring dengan perkembangan teknologi, metode berbasis kecerdasan buatan seperti machine learning menjadi alternatif yang menjanjikan dalam proses diagnosis medis. Salah satu pendekatan dalam machine learning yang dianggap efektif adalah ensemble learning, yang menggabungkan beberapa model untuk meningkatkan performa klasifikasi. Penelitian ini difokuskan pada penerapan tiga metode ensemble learning, yaitu Random Forest, AdaBoost, dan XGBoost, dalam klasifikasi kanker payudara berdasarkan dataset WBC.

2. METODE PENELITIAN

2.1 Pengumpulan Data

Dataset Data yang digunakan dalam penelitian ini adalah Wisconsin Breast Cancer (WBC) dari UCI Machine Learning Repository. Dataset ini terdiri dari 569 sampel dengan 30 fitur numerik yang berasal dari hasil digitalisasi citra biopsi sel kanker, serta satu label target yang menunjukkan kelas "Malignant" atau "Benign".

2.2 Pra Pemrosesan Data

Tahapan pra-pemrosesan data meliputi pengecekan nilai kosong, normalisasi fitur untuk menyamakan skala, pemisahan fitur (X) dan label (Y), serta pembagian data menggunakan metode 10-Fold Cross Validation untuk memperoleh hasil evaluasi yang lebih representatif.

2.3 Pengembangan Model

Tiga algoritma ensemble learning yang digunakan adalah Random Forest, AdaBoost, dan XGBoost. Masing-masing model dilatih menggunakan data latih hasil cross-validation dan diuji dengan data uji untuk memperoleh prediksi klasifikasi. Parameter dasar pada setiap model disesuaikan agar tidak menimbulkan overfitting dan tetap menjaga kecepatan komputasi.



Gambar 1 Grafik Alur Metode Pengembangan

2.4 Evaluasi Model

Evaluasi Model Evaluasi model Penedilakukan menggunakan beberapa metrik performa klasifikasi yaitu akurasi, presisi, recall, F1-score, AUC, dan confusion matrix. Evaluasi ini memberikan gambaran menyeluruh terhadap kekuatan dan kelemahan tiap model dalam mengklasifikasikan kanker payudara

Tabel 1 Hasil Matrik Evaluasi

Model	Accuracy	Precision	Recall	F1 Score
Random Forest	0.9649	0.9756	0.9302	0.9524
AdaBoost	0.9649	0.9756	0.9302	0.9524
XGBoost	0.9561	0.9524	0.9302	0.9412

Hasil eksperimen menunjukkan bahwa baik Random Forest maupun AdaBoost menghasilkan akurasi yang sama, yaitu sebesar 0.9649 atau 96.49%. Ini berarti bahwa hampir 96 dari setiap 100 data yang diuji dapat diklasifikasikan dengan benar oleh kedua algoritma tersebut. Nilai akurasi yang tinggi ini menandakan bahwa model memiliki kemampuan generalisasi yang baik terhadap data yang belum pernah dilihat sebelumnya. Di sisi lain, model XGBoost sedikit berada di bawah keduanya dengan akurasi sebesar 0.9561 atau 95.61%. Meskipun masih termasuk tinggi, perbedaan ini mengindikasikan bahwa XGBoost dalam kondisi default setting sedikit kurang optimal pada struktur dataset yang digunakan.

Model Random Forest menunjukkan performa yang sangat tinggi dengan akurasi sebesar

0.9649 atau setara dengan 96.49%, yang berarti bahwa dari seluruh data uji, sekitar 96 dari 100 prediksi dilakukan dengan benar. Hal ini menunjukkan bahwa model mampu membedakan dengan baik antara kasus kanker jinak dan ganas. Model AdaBoost menunjukkan hasil yang identik dengan Random Forest di seluruh metrik evaluasi, yaitu akurasi = 0.9649, presisi = 0.9756, recall = 0.9302, dan F1-score = 0.9524. Kesamaan ini menunjukkan bahwa algoritma boosting mampu mencapai tingkat akurasi dan presisi yang sama tinggi, meskipun mekanisme kerja algoritma berbeda. Presisi XGBoost sebesar 0.9524, lebih rendah dibandingkan dua model lainnya yang mencatat 0.9756. Ini berarti XGBoost lebih sering mengklasifikasikan kasus sebagai positif secara keliru (false positive) dibandingkan Random Forest dan AdaBoost. Meski demikian, recall XGBoost tetap sebesar 0.9302, yang menunjukkan bahwa model ini tetap konsisten dalam mendeteksi kanker ganas.

2.5 Cross Validation

Untuk mengukur konsistensi dan keandalan model dalam memprediksi diagnosis kanker payudara, dilakukan proses validasi silang sebanyak 10 kali lipat (10-fold cross-validation). Teknik ini membagi dataset menjadi sepuluh bagian yang proporsional, di mana sembilan bagian digunakan sebagai data pelatihan dan satu bagian sebagai data pengujian

Tabel 2 Pengukuran 10 Fold Cross Validation

Model	Akurasi	Presisi	Recall	F1-Score	AUC
Random Forest	0.9561	0.9491	0.9335	0.9398	0.9893
AdaBoost	0.9631	0.9780	0.9242	0.9480	0.9934
XGBoost	0.9491	0.9509	0.9100	0.9280	0.9893

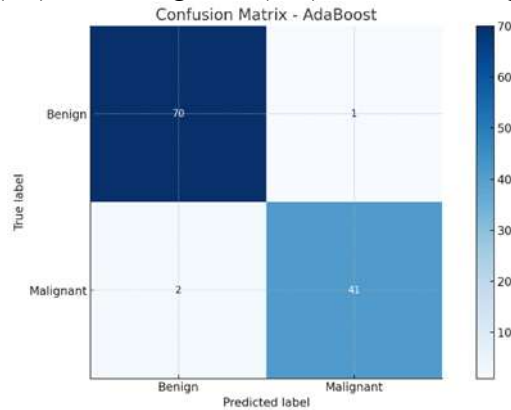
Model Random Forest menunjukkan performa klasifikasi yang sangat kompetitif, dengan rata-rata akurasi sebesar 95,61 persen. Nilai presisinya mencapai 94,91 persen, menandakan bahwa proporsi prediksi positif yang benar tergolong tinggi. Recall yang diperoleh sebesar 93,35 persen memperlihatkan bahwa model ini cukup andal dalam menangkap hampir seluruh kasus kanker positif yang ada dalam dataset. Keseimbangan antara presisi dan recall ditunjukkan oleh nilai F1-score sebesar 93,98 persen. Selain itu, Random Forest juga mencatatkan nilai AUC sebesar 0,9893, yang berarti model ini memiliki kemampuan diskriminasi yang sangat baik antara kelas kanker jinak dan ganas.

Berbeda dengan Random Forest, model AdaBoost memberikan hasil yang sedikit lebih unggul dalam beberapa aspek. Akurasi model ini tercatat sebesar 96,31 persen, menjadikannya sebagai model dengan nilai akurasi tertinggi di antara ketiganya. Presisi yang dihasilkan mencapai angka 97,80 persen, menunjukkan bahwa prediksi positif yang dilakukan oleh model hampir seluruhnya tepat sasaran. Walaupun nilai recall-nya sedikit lebih rendah dari Random Forest, yakni 92,42 persen, namun F1-score tetap tinggi pada angka 94,80 persen. Hal ini mengindikasikan bahwa AdaBoost mampu mempertahankan keseimbangan yang sangat baik antara kemampuan mendeteksi kasus positif dan menghindari prediksi palsu. Yang paling menonjol dari AdaBoost adalah nilai AUC yang mencapai 0,9934, menunjukkan performa diskriminatif terbaik dalam mengklasifikasikan dua kelas target secara konsisten.

Sementara itu, model XGBoost meskipun memiliki akurasi yang sedikit lebih rendah yakni sebesar 94,91 persen, tetap menampilkan performa yang solid. Presisi model ini berada di angka 95,09 persen, dan recall-nya sebesar 91,00 persen, mengindikasikan kemampuan yang cukup baik dalam mendeteksi kasus kanker positif, meski tidak setinggi dua model sebelumnya. Nilai F1-score yang diperoleh sebesar 92,80 persen menandakan bahwa meskipun model ini seimbang, performanya sedikit tertinggal dibandingkan Random Forest dan AdaBoost. Nilai AUC yang diraih adalah 0,9892, yang menunjukkan bahwa kemampuan model ini dalam membedakan antara dua kelas masih sangat baik, walaupun secara margin sedikit lebih rendah dari Random Forest dan AdaBoost.

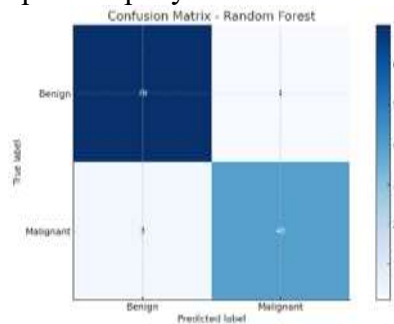
2.6 Confussion Matrix

Confusion matrix merupakan metode evaluasi yang digunakan untuk menilai performa klasifikasi dari suatu model, dengan menguraikan hasil prediksi ke dalam empat komponen utama, yaitu True Positive (TP), False Positive (FP), True Negative (TN), dan False Negative (FN).



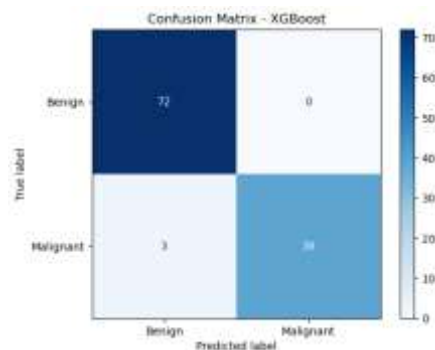
Gambar 2 Confusion Matrix – Adaboost

Model AdaBoost menunjukkan performa klasifikasi yang sangat baik terhadap data kanker payudara. Tingkat akurasi, presisi, dan recall yang tinggi menunjukkan bahwa model sangat efektif dalam membedakan antara kasus jinak dan ganas. Namun demikian, terdapat dua kasus False Negative yang perlu menjadi perhatian. Kesalahan prediksi pada kelas Malignant sebagai Benign berpotensi menyebabkan risiko serius dalam dunia nyata, khususnya dalam konteks medis yang menuntut sensitivitas tinggi terhadap kasus penyakit serius.



Gambar 3 Confusion Matrix – Random Forest

Model Random Forest menghasilkan total confusion matrix sebagai berikut: sebanyak 70 data pasien Benign berhasil diprediksi dengan benar (True Negative), sedangkan 40 data pasien Malignant juga diklasifikasikan dengan tepat (True Positive). Terdapat 1 data Benign yang salah diklasifikasikan sebagai Malignant (False Positive) dan 3 data Malignant yang keliru diklasifikasikan sebagai Benign (False Negative).



Gambar 4 Confusion Matrix – XGBoost

Berdasarkan confusion matrix model XGBoost yang ditampilkan, dapat dijelaskan bahwa

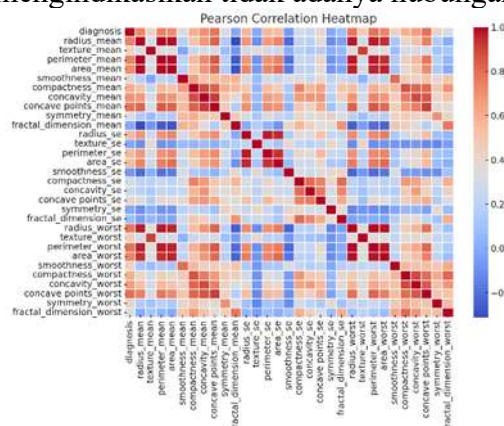
model ini menunjukkan performa yang sangat baik dalam mengenali data pasien dengan kondisi *Benign* secara akurat. Sebanyak 72 data yang tergolong *Benign* berhasil diklasifikasikan dengan benar oleh model, tanpa adanya kesalahan klasifikasi ke dalam kelas *Malignant*. Hal ini mengindikasikan bahwa nilai *False Positive* adalah nol, yang berarti model tidak memberikan peringatan palsu terhadap pasien yang sehat.

Sementara itu, untuk data pasien dengan kondisi *Malignant*, sebanyak 39 data berhasil dikenali dengan tepat sebagai *Malignant*, yang menunjukkan kemampuan model dalam melakukan deteksi terhadap kasus kanker. Namun, masih terdapat 3 data pasien dengan label aktual *Malignant* yang secara keliru diklasifikasikan sebagai *Benign*. Kesalahan ini tergolong *False Negative*, yaitu kondisi di mana pasien yang sebenarnya mengidap kanker tidak berhasil terdeteksi oleh model.

Kondisi ini menandakan bahwa meskipun XGBoost mampu menghindari kesalahan dalam mengidentifikasi pasien sehat (bebas kanker), namun masih terdapat tantangan dalam meningkatkan sensitivitas terhadap kasus kanker yang sebenarnya. Oleh karena itu, meskipun secara keseluruhan performanya kuat, model XGBoost tetap perlu dilakukan penyempurnaan untuk mengurangi potensi kesalahan *False Negative* yang dapat berdampak signifikan dalam konteks medis.

2.7 Pearson Correlation

Analisis korelasi pada dataset Wisconsin Breast Cancer dilakukan untuk mengetahui hubungan linear antara setiap fitur dengan variabel target, yaitu diagnosis, serta untuk memahami interkoneksi antar fitur yang ada. Korelasi dihitung menggunakan koefisien Pearson yang memiliki rentang nilai antara -1 hingga +1, di mana nilai mendekati +1 menunjukkan hubungan linear positif yang sangat kuat, nilai mendekati -1 menunjukkan hubungan linear negatif yang sangat kuat, sedangkan nilai mendekati 0 mengindikasikan tidak adanya hubungan linear yang signifikan.



Gambar 4 Grafik Pearson Correlation

Hasil analisis menunjukkan bahwa sebagian besar fitur pada dataset memiliki korelasi positif terhadap variabel diagnosis. Fitur-fitur geometrik seperti *radius_mean*, *perimeter_mean*, dan *area_mean* memperlihatkan nilai korelasi yang tinggi terhadap diagnosis, masing-masing dengan nilai koefisien Pearson sebesar 0,73, 0,74, dan 0,71. Hal ini mengindikasikan bahwa semakin besar ukuran rata-rata radius, perimeter, dan area dari sel kanker yang dianalisis, semakin besar kemungkinan tumor tersebut bersifat malignan. Selain itu, fitur yang menggambarkan ketidakrataan permukaan sel, yaitu *concavity_mean* dan *concave points_mean*, juga menunjukkan korelasi yang tinggi terhadap diagnosis dengan nilai koefisien sebesar 0,70 dan 0,77. Fitur-fitur ini mencerminkan karakteristik morfologi yang lebih agresif dan tidak teratur yang sering ditemukan pada sel kanker ganas.

3. HASIL DAN PEMBAHASAN

Hasil evaluasi menunjukkan bahwa ketiga model memiliki performa klasifikasi yang baik, dengan akurasi di atas 94%. Model AdaBoost memperoleh akurasi tertinggi sebesar 96,31%, diikuti oleh Random Forest sebesar 95,61%, dan XGBoost sebesar 94,91%. Dari confusion matrix,

diketahui bahwa AdaBoost memiliki distribusi klasifikasi yang paling seimbang dengan jumlah kesalahan klasifikasi paling sedikit. Model XGBoost menunjukkan presisi tinggi tanpa False Positive, namun masih terdapat tiga False Negative. Sedangkan Random Forest memberikan hasil yang stabil dengan satu False Positive dan tiga False Negative.

Hasil metrik evaluasi yang dihitung dari confusion matrix ini memperkuat penilaian terhadap kinerja model. Tingkat akurasi keseluruhan diperkirakan sebesar 97,37%, yang menunjukkan bahwa dari seluruh data uji, hampir semua sampel berhasil diklasifikasikan dengan benar. Precision pada kelas malignant diperkirakan mencapai 100%, yang berarti semua prediksi positif model adalah benar-benar kasus malignant tanpa ada kesalahan prediksi positif palsu. Namun, nilai recall diperkirakan sedikit lebih rendah karena keberadaan tiga kasus False Negative, yang menunjukkan bahwa model gagal mendeteksi sebagian kecil kanker ganas. Nilai F1-Score yang mendekati optimal memperlihatkan keseimbangan yang baik antara precision dan recall, meskipun ada sedikit trade-off pada sensitivitas.

4. KESIMPULAN

Penelitian ini berhasil membuat metode evaluasi yang menunjukkan bahwa ketiga metode ensemble yang digunakan memiliki performa yang tinggi, dengan akurasi keseluruhan di atas 94%. Model Random Forest mencatat akurasi sebesar 95,61%, AdaBoost sebesar 96,31%, dan XGBoost sebesar 94,91%. Dari analisis confusion matrix, diketahui bahwa AdaBoost menunjukkan performa yang paling seimbang, dengan hanya satu kesalahan False Positive dan dua False Negative. XGBoost menunjukkan presisi sempurna tanpa False Positive, namun menghasilkan tiga False Negative. Sedangkan Random Forest memberikan distribusi hasil yang relatif stabil, tetapi masih terdapat satu False Positive dan tiga False Negative.

Disini dapat disimpulkan bahwa algoritma ensemble learning dapat secara efektif digunakan dalam klasifikasi kanker payudara, dengan model AdaBoost menunjukkan performa terbaik secara keseluruhan. Hasil ini menunjukkan bahwa pemanfaatan machine learning dapat menjadi alternatif solusi dalam sistem deteksi dini kanker, serta mendukung proses pengambilan keputusan medis. Hal tersebut karena kemampuannya dalam menjaga keseimbangan antara deteksi terhadap pasien kanker dan minimnya kesalahan klasifikasi dibandingkan dengan dua model lainnya. Temuan ini menunjukkan bahwa AdaBoost merupakan metode yang menjanjikan dan potensial untuk dikembangkan lebih lanjut dalam sistem deteksi dini kanker payudara berbasis machine learning. Pengembangan lebih lanjut dengan optimasi parameter dan penerapan pada data klinis nyata sangat disarankan untuk meningkatkan akurasi dan keandalan sistem.

DAFTAR PUSTAKA

- [1] Agusviyanda Agusviyanda, R. N. (2025). PENINGKATAN ALGORITMA C4.5 MENGGUNAKAN ENSEMBLE LEARNING UNTUK MENDETEKSI PENYAKIT GINJAL. doi:10.36987/informatika
- [2] Chaofeng Yuan, N. W. (2023, January). Rebalance Weights AdaBoost-SVM Model for Imbalanced Data. *Computational Intelligence and Neuroscience*, 1-26. doi:10.1155/2023/4860536
- [3] Deni Kurnia, M. I. (2023). Seleksi Fitur dengan Particle Swarm Optimization pada Klasifikasi Penyakit Parkinson Menggunakan XGBoost. *10*. doi:https://doi.org/10.25126/jtiik.2023107252
- [4] Foundation, B. C. (2021). *bcrf.org*. Retrieved from <https://www.bcrf.org/breast-cancer-statistics-and-resources/?utm>
- [5] Idris Nur Farahaina, I. M. (2024). Tinjauan metode ensemble homogen pada klasifikasi data kanker payudara. *Tinjauan Teknik Elektro*, 101 - 104. doi:10.15199/48.2024.01.21
- [6] Jafari, A. (2023). Machine-learning methods in detecting breast cancer and related therapeutic issues: a review. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 12(1). doi:https://doi.org/10.1080/21681163.2023.2299093

- [7] Kemenkes RI. (2021). Retrieved from Kemenkes Unit Pelayanan Kesehatan: <https://upk.kemkes.go.id/new/deteksi-dini-kanker-payudara-dengan-sadari-dan-sadanis>
- [8] Lucky Lhaura Van FC, M. K. (2025). The Development of Stacking Techniques in Machine Learning for Breast Cancer Detection. *Journal Of Applied Data Science*. doi:<https://doi.org/10.47738/jads.v6i1.416>
- [9] Maulidina Cahaya Rani, A. S. (2025). Perbandingan Algoritma Random Forest, Naive Bayes, Dan Neural Network Dalam Klasifikasi Penyakit Jantung. *Jurnal Sains Informatika Terapan*, 4. Retrieved from <https://rcf-indonesia.org/home/index.php/jsit/article/view/609>
- [10] MUHAMAD FADLI, R. A. (2023). KLASIFIKASI DAN EVALUASI PERFORMA MODEL RANDOM FOREST UNTUK PREDIKSI STROKE. 12. doi:<http://dx.doi.org/10.31000/jt.v12i2.9099>
- [11] Nawangsih, A. M. (2024). Perbandingan Metode Ensemble Untuk Meningkatkan Akurasi Algoritma Machine Learning Dalam Memprediksi Penyakit Breast Cancer (Kanker Payudara). *Jurnal Informatika dan Sistem Informasi*. Retrieved from <https://ojs.trigunadharma.ac.id/index.php/jis/article/view/9563>
- [12] Nur Kholis Majid, C. S. (2025). Peningkatan Keberagaman Data untuk Klasifikasi Penyakit Diabetes Berbasis Stacking Ensemble Learning. 10. Retrieved from <https://ejournal.poltekharber.ac.id/index.php/informatika/article/view/7375>
- [13] Organization, W. H. (2022). *The Global Breast Cancer Initiative*. Retrieved from <https://www.who.int/initiatives/global-breast-cancer-initiative?utm>
- [14] Setyawan Wibisono, W. H. (2024). Optimalisasi Model Klasifikasi Diabetes Menggunakan Ensemble Learning Adaboost, Gradient Boosting, dan XGBoost. *Journal of Data Science and Technology*, 45-55. Retrieved from https://scholar.google.com/citations?view_op=view_citation&hl=id&user=44ei0EUAAAAJ&sortby=pubdate&citation_for_view=44ei0EUAAAAJ:u9iWguZQMMsC
- [15] Syed Mahfuz Al Hasan, D. L. (2025). Screening programmes and breast cancer mortality: an. *Screening programmes and breast cancer mortality*. Retrieved from https://cdn.who.int/media/docs/default-source/bulletin/online-first/blt.24.292529.pdf?sfvrsn=2c78d36f_3&utm
- [16] WHO), W. H. (2024, February). *Global Cancer Observatory - World Fact Sheet*. Retrieved from <https://gco.iarc.who.int/media/globocan/factsheets/populations/900-world-fact-sheet.pdf>