

**ANALISIS PERBANDINGAN MACHINE LEARNING
DENGAN HYPERPARAMETER TUNING TERHADAP
PENYAKIT KANKER PAYUDARA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

NOVANDA AFARIZ YUDI BATARA

21.11.4152

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS PERBANDINGAN MACHINE LEARNING
DENGAN HYPERPARAMETER TUNING TERHADAP
PENYAKIT KANKER PAYUDARA**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

NOVANDA AFARIZ YUDI BATARA

21.11.4152

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS PERBANDINGAN MACHINE LEARNING DENGAN
HYPERPARAMETER TUNING TERHADAP PENYAKIT KANKER
PAYUDARA**

yang disusun dan diajukan oleh

Novanda Afariz Yudi Batara

21.11.4152

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Mei 2025

Dosen Pembimbing,



Kusniwi, S.Kom., M.Eng.
NIK. 190302112

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS PERBANDINGAN MACHINE LEARNING DENGAN
HYPERPARAMETER TUNING TERHADAP PENYAKIT KANKER
PAYUDARA

yang disusun dan diajukan oleh

Novanda Afariz Yudi Batara

21.11.4152

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Mei 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ainul Yaqin, S.Kom., M.Kom.
NIK. 190302255

Anna Baita, S.Kom., M.Kom.
NIK. 190302290

Kusnawi, S.Kom., M.Eng.
NIK. 190302112



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Mei 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Novanda Afariz Yudi Batara

NIM : 21.11.4152

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Perbandingan Machine Learning Dengan Hyperparameter Tuning Terhadap Penyakit Kanker Payudara

Dosen Pembimbing : Kusnawi, S.Kom., M.Eng.

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas **AMIKOM** Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas **AMIKOM** Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 23 Mei 2025

Yang Menyatakan,



Novanda Afariz Yudi Batara

HALAMAN PERSEMBAHAN

Dengan segala kerendahan hati dan rasa syukur yang mendalam, skripsi ini saya persembahkan kepada:

Tuhan Yang Maha Esa, atas limpahan rahmat, petunjuk, dan kekuatan yang senantiasa menyertai setiap langkah dalam proses kehidupan dan penyusunan karya ilmiah ini.

Kedua orang tua saya tercinta, yang dengan tulus memberikan kasih sayang, doa, dan dukungan tanpa batas. Semangat dan keteguhan hati kalian menjadi landasan utama dalam pencapaian ini.

Adik saya, Charolina Riska Maharani, yang selalu menjadi sumber semangat dan inspirasi dalam menjalani setiap proses.

Serta Atika Rifanandasari, support system yang setia mendampingi dan memberi kekuatan di setiap tantangan, terima kasih atas kehadiran, doa, dan motivasi yang tak ternilai.

Semoga karya ini menjadi bentuk kecil dari rasa terima kasih atas segala cinta dan dukungan yang telah saya terima.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika.

Penyusunan skripsi ini tidak terlepas dari bimbingan, dukungan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih yang sebesar-besarnya kepada Bapak Kusnawi, S.Kom., M.Eng., selaku dosen pembimbing yang telah dengan sabar membimbing dan memberikan arahan selama proses penyusunan skripsi ini. Ucapan terima kasih juga penulis sampaikan kepada tim dosen penguji yang telah memberikan masukan dan koreksi yang sangat berharga demi penyempurnaan karya ini.

Penulis juga menyampaikan apresiasi yang mendalam kepada seluruh dosen dan staf di Program Studi Informatika atas ilmu, bantuan, dan pelayanan yang telah diberikan selama masa studi. Ucapan terima kasih yang tulus penulis sampaikan kepada kedua orang tua tercinta atas doa, kasih sayang, serta dukungan moril dan materiil yang tidak pernah terputus sepanjang perjalanan studi ini. Tidak lupa, penulis juga mengucapkan terima kasih kepada pihak-pihak lain yang turut membantu dan memberikan semangat dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis terbuka terhadap segala saran dan kritik yang bersifat membangun. Semoga skripsi ini dapat memberikan manfaat dan kontribusi positif bagi pengembangan ilmu pengetahuan dan bagi semua pihak yang membacanya.

Yogyakarta, 23 Mei 2025

Novanda Afariz Yudi Batara

INTISARI

Kanker payudara merupakan salah satu penyakit penyebab kematian tertinggi di dunia, khususnya pada wanita. Deteksi dini sangat penting untuk meningkatkan peluang kesembuhan, namun tantangan tetap ada dalam mengklasifikasikan data medis dengan akurasi tinggi. Penelitian ini bertujuan untuk membandingkan performa empat algoritma machine learning Support Vector Machine (SVM), Random Forest, K-Nearest Neighbor (KNN), dan Gradient Boosting dalam klasifikasi kanker payudara, dengan dukungan proses hyperparameter tuning menggunakan metode GridSearchCV dan RandomizedSearchCV. Dataset yang digunakan berasal dari Kaggle, terdiri atas 4024 baris dan 16 fitur, yang kemudian melalui tahapan preprocessing, normalisasi, SMOTE, dan pembagian data dengan K-Fold Cross Validation. Penelitian ini mengimplementasikan pemodelan dan evaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma Random Forest dengan tuning menghasilkan performa terbaik, diikuti oleh Gradient Boosting, KNN, dan SVM. Penerapan hyperparameter tuning terbukti mampu meningkatkan akurasi model secara signifikan dibandingkan sebelum tuning. Kontribusi penelitian ini memberikan referensi metode terbaik dalam klasifikasi data kanker payudara dan dapat dimanfaatkan oleh praktisi data dan tenaga medis untuk mendukung sistem deteksi dini berbasis teknologi machine learning. Penelitian selanjutnya disarankan untuk mengintegrasikan teknik feature engineering lanjutan dan algoritma deep learning.

Kata kunci: kanker payudara, klasifikasi, machine learning, hyperparameter tuning, GridSearchCV

ABSTRACT

Breast cancer is one of the leading causes of death in the world, especially in women. Early detection is crucial to increase the chance of cure, but challenges remain in classifying medical data with high accuracy. This study aims to compare the performance of four machine learning algorithms Support Vector Machine (SVM), Random Forest, K-Nearest Neighbor (KNN), and Gradient Boosting in breast cancer classification, with the support of hyperparameter tuning process using GridSearchCV and RandomizedSearchCV methods. The dataset used comes from Kaggle, consisting of 4024 rows and 16 features, which then went through preprocessing, normalization, SMOTE, and data sharing with K-Fold Cross Validation. This research implements modeling and evaluation using accuracy, precision, recall, and F1-score metrics. The results show that Random Forest algorithm with tuning produces the best performance, followed by Gradient Boosting, KNN, and SVM. The application of hyperparameter tuning is proven to be able to significantly improve the accuracy of the model compared to before tuning. The contribution of this research provides a reference to the best method in breast cancer data classification and can be utilized by data practitioners and medical personnel to support early detection systems based on machine learning technology. Future research is recommended to integrate advanced feature engineering techniques and deep learning algorithms.

Keywords: breast cancer, classification, machine learning, hyperparameter tuning, GridSearchCV