

---

## Analisis Kombinasi *Bagging* dan *Gradient Boosting* pada Klasifikasi Diagnosis Kanker Payudara

Rully Pramudita<sup>1\*</sup>, Dwi Ismiyana Putri<sup>2</sup>, Bambang Kriswantara<sup>3</sup>, Vina Zahrotun Nazah<sup>4</sup>  
<sup>1,4</sup> Teknik Informatika/Universitas Bina Insani. <sup>2</sup>Sistem Informasi/Universitas Bina Insani. <sup>3</sup>Manajemen  
Informatika/Universitas Bina Insani.

<sup>1</sup>Email<sup>1</sup>: [rullypramudita@binainsani.ac.id](mailto:rullypramudita@binainsani.ac.id)

<sup>2</sup>Email<sup>2</sup>: [dwiismiyapanutri@binainsani.ac.id](mailto:dwiismiyapanutri@binainsani.ac.id)

<sup>3</sup>Email<sup>3</sup>: [bambangkris@binainsani.ac.id](mailto:bambangkris@binainsani.ac.id)

<sup>4</sup>Email<sup>4</sup>: [vinazahrotunnazah@binainsani.ac.id](mailto:vinazahrotunnazah@binainsani.ac.id)

\*) Corresponding Author

---

### ABSTRACT

*This study aims to analyze the performance of Gradient Boosting in improving the classification accuracy of several machine learning algorithms for breast cancer diagnosis. The dataset used in this study is the Breast Cancer Wisconsin (Diagnostic) dataset consisting of 569 records with 30 numerical attributes representing cell characteristics, and two diagnosis classes: malignant and benign. The experimental process was conducted in several stages, starting with training baseline classification models using five algorithms, namely Support Vector Machine (SVM), Neural Network (NN), Logistic Regression (LR), Decision Tree (DT), and K-Nearest Neighbors (KNN). After that, the models were optimized using the Gradient Boosting technique to evaluate its impact on model performance. The dataset was divided into 70% training data and 30% testing data to ensure objective evaluation. Model performance was assessed using confusion matrix, Receiver Operating Characteristic (ROC) curve, and Area Under Curve (AUC). The results show that ensemble optimization using boosting techniques can improve the classification capability of several algorithms. Neural Network achieved the highest baseline accuracy of 98.2%, while Logistic Regression produced the most stable performance with the highest AUC value of 99.7%. These findings indicate that Gradient Boosting can enhance the predictive performance of machine learning models in medical classification tasks, particularly in breast cancer diagnosis.*

**Keywords:** Machine Learning; Gradient Boosting; Breast Cancer Classification; Ensemble Learning; AUC.

### ABSTRAK

*Penelitian ini bertujuan untuk menganalisis kinerja metode Gradient Boosting dalam meningkatkan performa beberapa algoritma klasifikasi machine learning pada diagnosis kanker payudara. Dataset yang digunakan adalah Breast Cancer Wisconsin (Diagnostic) yang terdiri dari 569 data dengan 30 atribut numerik yang merepresentasikan karakteristik sel jaringan payudara serta dua kelas diagnosis yaitu malignant dan benign. Proses eksperimen dilakukan melalui beberapa tahap, dimulai dari pelatihan model klasifikasi dasar menggunakan lima algoritma yaitu Support Vector Machine (SVM), Neural Network (NN), Logistic Regression (LR), Decision Tree (DT), dan K-Nearest Neighbors (KNN). Selanjutnya model tersebut dioptimasi menggunakan teknik Gradient Boosting untuk melihat pengaruhnya terhadap peningkatan performa model. Dataset dibagi menjadi 70% data training dan 30% data testing untuk memastikan evaluasi yang objektif. Evaluasi kinerja model dilakukan menggunakan confusion matrix, kurva Receiver Operating Characteristic (ROC), dan Area Under Curve (AUC). Hasil penelitian menunjukkan bahwa optimasi menggunakan teknik boosting mampu meningkatkan kemampuan klasifikasi pada beberapa algoritma machine learning. Neural Network memperoleh akurasi dasar tertinggi sebesar 98,2%, sedangkan Logistic Regression menunjukkan performa paling stabil dengan nilai AUC tertinggi sebesar 99,7%. Hasil ini menunjukkan*

---

*bahwa metode Gradient Boosting dapat meningkatkan kemampuan prediksi model machine learning pada tugas klasifikasi medis khususnya diagnosis kanker payudara.*

**Keywords:** *Machine Learning, Gradient Boosting, Klasifikasi Kanker Payudara, Ensemble Learning, AUC.*

---

## A. PENDAHULUAN

Perkembangan teknologi Machine Learning (ML) telah memberikan kontribusi signifikan dalam berbagai bidang yang memanfaatkan data sebagai dasar pengambilan keputusan, termasuk pada bidang kesehatan [1]. Salah satu tugas utama dalam Machine Learning adalah klasifikasi, yaitu proses mengelompokkan data ke dalam kategori tertentu berdasarkan karakteristik atau fitur yang dimiliki. Metode klasifikasi telah banyak diterapkan pada berbagai permasalahan seperti deteksi penyakit, prediksi keuangan, analisis perilaku pengguna, hingga sistem rekomendasi. Dalam konteks medis, kemampuan model klasifikasi dalam mengenali pola dari data klinis dapat membantu tenaga kesehatan dalam melakukan diagnosis penyakit secara lebih cepat dan akurat [2], [3], [4].

Salah satu permasalahan yang banyak diteliti dalam bidang klasifikasi medis adalah diagnosis kanker payudara. Beberapa penelitian menunjukkan bahwa penerapan teknik machine learning mampu meningkatkan kemampuan analisis data medis serta membantu proses pengambilan keputusan dalam sistem diagnosis penyakit. Pemanfaatan metode klasifikasi pada data medis memungkinkan model pembelajaran mesin untuk mengenali pola yang kompleks pada karakteristik sel jaringan sehingga dapat membedakan antara kondisi malignant dan benign secara lebih akurat [5]–[7].

Berbagai algoritma klasifikasi telah digunakan untuk memprediksi diagnosis kanker payudara, di antaranya Support Vector Machine (SVM), Neural Network (NN), Logistic Regression (LR), Decision Tree (DT), dan K-Nearest Neighbors (KNN). Setiap algoritma memiliki pendekatan yang berbeda dalam mempelajari pola data sehingga menghasilkan tingkat performa yang bervariasi tergantung pada karakteristik dataset yang digunakan [2], [3], [8], [9], [10], [11]. Namun demikian, penggunaan model tunggal seringkali memiliki keterbatasan dalam hal stabilitas model dan kemampuan generalisasi terhadap data baru.

Untuk mengatasi keterbatasan tersebut, pendekatan ensemble learning banyak digunakan dalam penelitian Machine Learning. Ensemble learning merupakan teknik yang menggabungkan beberapa model pembelajaran untuk menghasilkan performa prediksi yang lebih baik dibandingkan dengan model tunggal [12]. Metode ini bekerja dengan memanfaatkan kekuatan beberapa model sehingga mampu mengurangi variansi dan meningkatkan stabilitas model klasifikasi.

Salah satu pendekatan ensemble yang banyak digunakan adalah metode boosting. Konsep boosting pertama kali dikembangkan melalui pendekatan gradient boosting yang bekerja dengan membangun model secara bertahap untuk memperbaiki kesalahan prediksi pada model sebelumnya [13]. Pendekatan ini memungkinkan model untuk meminimalkan fungsi kesalahan secara iteratif sehingga mampu meningkatkan kemampuan prediksi pada dataset yang kompleks.

Perkembangan lebih lanjut dari teknik boosting juga menghasilkan berbagai varian algoritma seperti Extreme Gradient Boosting (XGBoost) yang dikenal memiliki performa tinggi dan efisiensi komputasi yang baik dalam berbagai kasus klasifikasi dan prediksi [7]. Selain itu, beberapa penelitian juga menunjukkan bahwa penerapan teknik ensemble learning dapat meningkatkan performa model klasifikasi pada berbagai domain aplikasi seperti prediksi risiko kesehatan, analisis data medis, serta klasifikasi berbasis machine learning [9]–[11], [14].

Penelitian sebelumnya juga menunjukkan bahwa penerapan teknik ensemble seperti Bagging dan Boosting mampu meningkatkan stabilitas serta akurasi model klasifikasi pada berbagai kasus

prediksi berbasis data. Teknik Bagging bekerja dengan membuat beberapa subset data menggunakan metode bootstrap sampling untuk mengurangi variansi model, sedangkan Boosting berfokus pada peningkatan performa model dengan memperbaiki kesalahan klasifikasi secara bertahap [5], [15].

Berdasarkan kajian literatur tersebut, penelitian ini bertujuan untuk menganalisis pengaruh penerapan metode Gradient Boosting terhadap performa beberapa algoritma klasifikasi Machine Learning dalam kasus diagnosis kanker payudara. Lima algoritma klasifikasi digunakan sebagai model dasar yaitu Support Vector Machine, Neural Network, Logistic Regression, Decision Tree, dan K-Nearest Neighbors. Setiap model dianalisis performanya sebelum dan sesudah penerapan metode Gradient Boosting untuk mengetahui sejauh mana teknik ensemble tersebut mampu meningkatkan kinerja model klasifikasi.

Dengan menggunakan metrik evaluasi seperti confusion matrix, kurva Receiver Operating Characteristic (ROC), dan Area Under Curve (AUC), penelitian ini diharapkan dapat memberikan gambaran yang lebih komprehensif mengenai efektivitas metode Gradient Boosting dalam meningkatkan performa algoritma klasifikasi pada diagnosis kanker payudara.

## B. METODE

Penelitian ini menggunakan pendekatan eksperimen komputasional untuk menganalisis pengaruh penerapan metode ensemble learning terhadap performa beberapa algoritma klasifikasi pada kasus diagnosis kanker payudara. Pendekatan ini dilakukan dengan membangun model klasifikasi menggunakan beberapa algoritma machine learning, kemudian mengevaluasi peningkatan performa model setelah diterapkan teknik optimasi berbasis ensemble learning.

### 1. Dataset

Dataset yang digunakan dalam penelitian ini adalah Breast Cancer Wisconsin (Diagnostic) yang tersedia secara publik melalui platform Kaggle dan juga dapat diakses melalui library sklearn. Dataset ini terdiri dari 569 data observasi dengan 30 atribut numerik yang merepresentasikan karakteristik inti sel jaringan payudara. Dataset ini memiliki dua kelas diagnosis yaitu malignant (ganas) dan benign (jinak). Seluruh atribut pada dataset tidak memiliki nilai yang hilang sehingga dapat langsung digunakan pada proses pemodelan.

Sebelum proses pemodelan dilakukan, dataset dibagi menjadi dua bagian yaitu data training sebesar 70% dan data testing sebesar 30%. Pembagian ini bertujuan untuk memastikan bahwa model yang dibangun dapat dievaluasi menggunakan data yang tidak digunakan pada proses pelatihan sehingga hasil evaluasi lebih objektif.

### 2. Tahapan Penelitian

Tahapan penelitian terdiri dari beberapa proses utama yaitu *data preprocessing*, pelatihan model, optimasi model menggunakan metode ensemble learning, serta evaluasi performa model. Alur proses penelitian ditunjukkan pada Gambar 1.

Pada tahap awal dilakukan proses data preprocessing yang meliputi pengambilan dataset dan pembagian data menjadi data training dan data testing. Data training digunakan untuk melatih model klasifikasi sedangkan data testing digunakan untuk mengevaluasi performa model.

Tahap berikutnya adalah pelatihan model klasifikasi dasar menggunakan lima algoritma machine learning yaitu:

- a. Support Vector Machine (SVM)
- b. Neural Network (NN)
- c. Logistic Regression (LR)
- d. Decision Tree (DT)

e. K-Nearest Neighbors (KNN)

Kelima algoritma tersebut dipilih karena memiliki karakteristik yang berbeda dalam mempelajari pola data sehingga memungkinkan dilakukan analisis perbandingan performa model secara lebih komprehensif.

Setelah model dasar dilatih, proses berikutnya adalah penerapan metode Bagging (*Bootstrap Aggregating*). Pada tahap ini data training dibagi menjadi beberapa subset melalui teknik bootstrap sampling. Setiap subset data digunakan untuk melatih model klasifikasi sehingga dihasilkan beberapa model yang kemudian digabungkan untuk meningkatkan stabilitas dan mengurangi variansi model.

Selanjutnya dilakukan proses optimasi menggunakan metode *Gradient Boosting*. Metode ini bekerja dengan membangun model prediksi secara bertahap dimana setiap model baru berfungsi untuk memperbaiki kesalahan prediksi yang dihasilkan oleh model sebelumnya. Dengan pendekatan ini, model secara iteratif mampu meminimalkan kesalahan prediksi dan meningkatkan performa klasifikasi.

Pada penelitian ini, metode *Gradient Boosting* diterapkan pada hasil model yang telah melalui proses Bagging. Dengan demikian, setiap algoritma klasifikasi dasar akan memiliki model hasil optimasi menggunakan kombinasi teknik Bagging dan *Gradient Boosting*.

### 3. Evaluasi Model

Evaluasi model dilakukan untuk mengetahui performa setiap model klasifikasi baik model dasar maupun model hasil optimasi ensemble learning. Evaluasi dilakukan menggunakan beberapa metrik pengukuran yaitu:

a. Confusion Matrix

Confusion matrix digunakan untuk melihat jumlah prediksi yang benar maupun salah dari setiap kelas sehingga dapat dihitung nilai akurasi model.

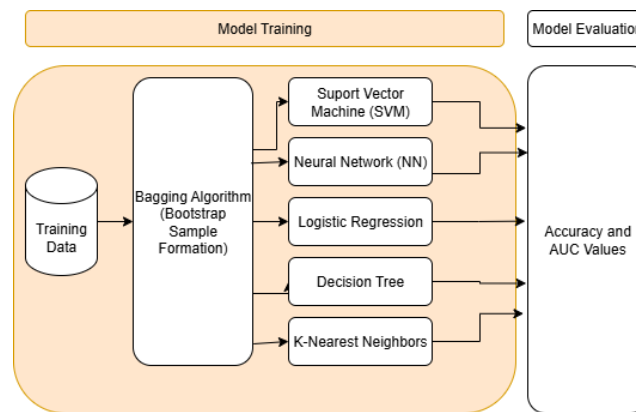
b. Receiver Operating Characteristic (ROC)

Kurva ROC digunakan untuk menggambarkan hubungan antara *true positive rate* dan *false positive rate* pada berbagai nilai ambang batas klasifikasi.

c. Area Under Curve (AUC)

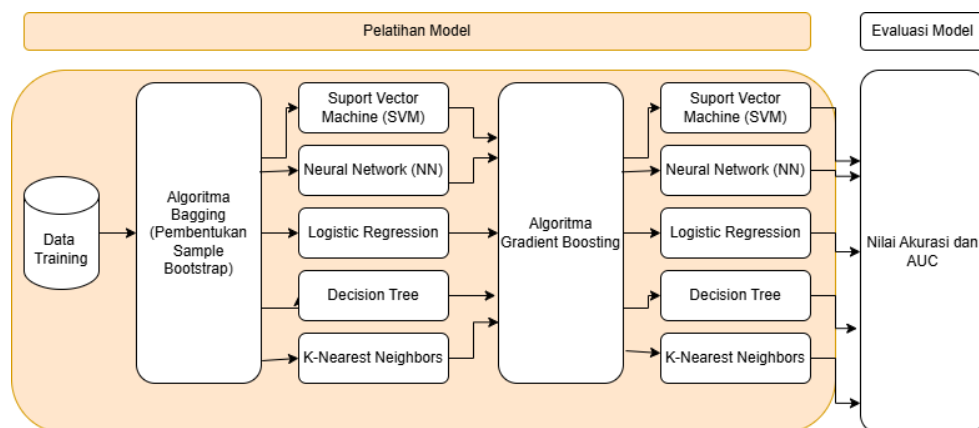
AUC digunakan sebagai ukuran numerik yang menunjukkan kemampuan model dalam membedakan antara kelas positif dan negatif.

Hasil evaluasi dari setiap model kemudian dibandingkan untuk mengetahui sejauh mana penerapan metode ensemble learning khususnya Gradient Boosting mampu meningkatkan performa algoritma klasifikasi pada diagnosis kanker payudara.



Gambar 1. Proses Pelatihan Model Menggunakan Algoritma Bagging

Gambar 1 menunjukkan proses pelatihan model dimulai dari data training yang diproses menggunakan algoritma Bagging untuk membentuk beberapa sampel bootstrap. Sampel tersebut kemudian digunakan untuk melatih lima algoritma klasifikasi yaitu SVM, Neural Network, Logistic Regression, Decision Tree, dan K-Nearest Neighbors. Hasil dari setiap model selanjutnya dievaluasi menggunakan metrik akurasi dan AUC.



Gambar 2. Proses Pelatihan Model Menggunakan Bagging dan Gradient Boosting

Gambar 2 menunjukkan alur pengembangan model dimulai dari data training yang diproses menggunakan algoritma Bagging untuk membentuk sampel bootstrap. Hasil model dari setiap algoritma klasifikasi kemudian diproses kembali menggunakan algoritma Gradient Boosting untuk meningkatkan performa model. Model akhir dari masing-masing algoritma kemudian dievaluasi menggunakan nilai akurasi dan AUC.

## C. HASIL DAN PEMBAHASAN

### 1. HASIL

Pengujian model dilakukan untuk mengetahui pengaruh penerapan teknik ensemble learning terhadap performa algoritma klasifikasi pada diagnosis kanker payudara. Pengujian dilakukan secara bertahap mulai dari model dasar, model yang dioptimasi menggunakan metode Bagging, hingga model yang dioptimasi menggunakan kombinasi Bagging dan Gradient Boosting.

Dataset yang digunakan pada penelitian ini adalah Breast Cancer Wisconsin (Diagnostic) yang terdiri dari 569 data dengan 30 atribut numerik serta dua kelas diagnosis yaitu malignant dan benign. Data dibagi menjadi 70% data training dan 30% data testing untuk memastikan evaluasi model dilakukan secara objektif.

#### a. Hasil Akurasi Model

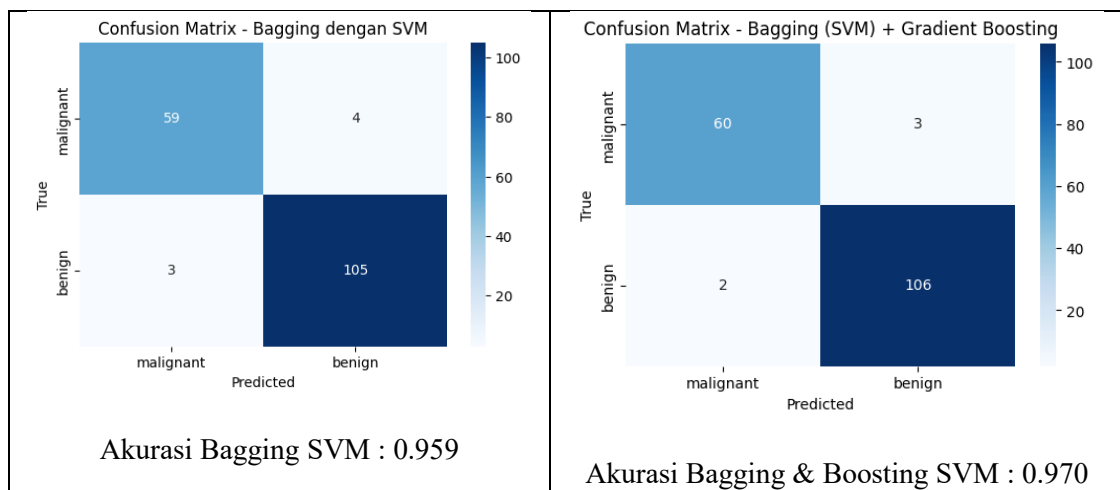
Tabel 1 menunjukkan perbandingan nilai akurasi antara model yang menggunakan metode Bagging dan model yang menggunakan kombinasi Bagging dan Gradient Boosting berdasarkan hasil eksperimen penelitian.

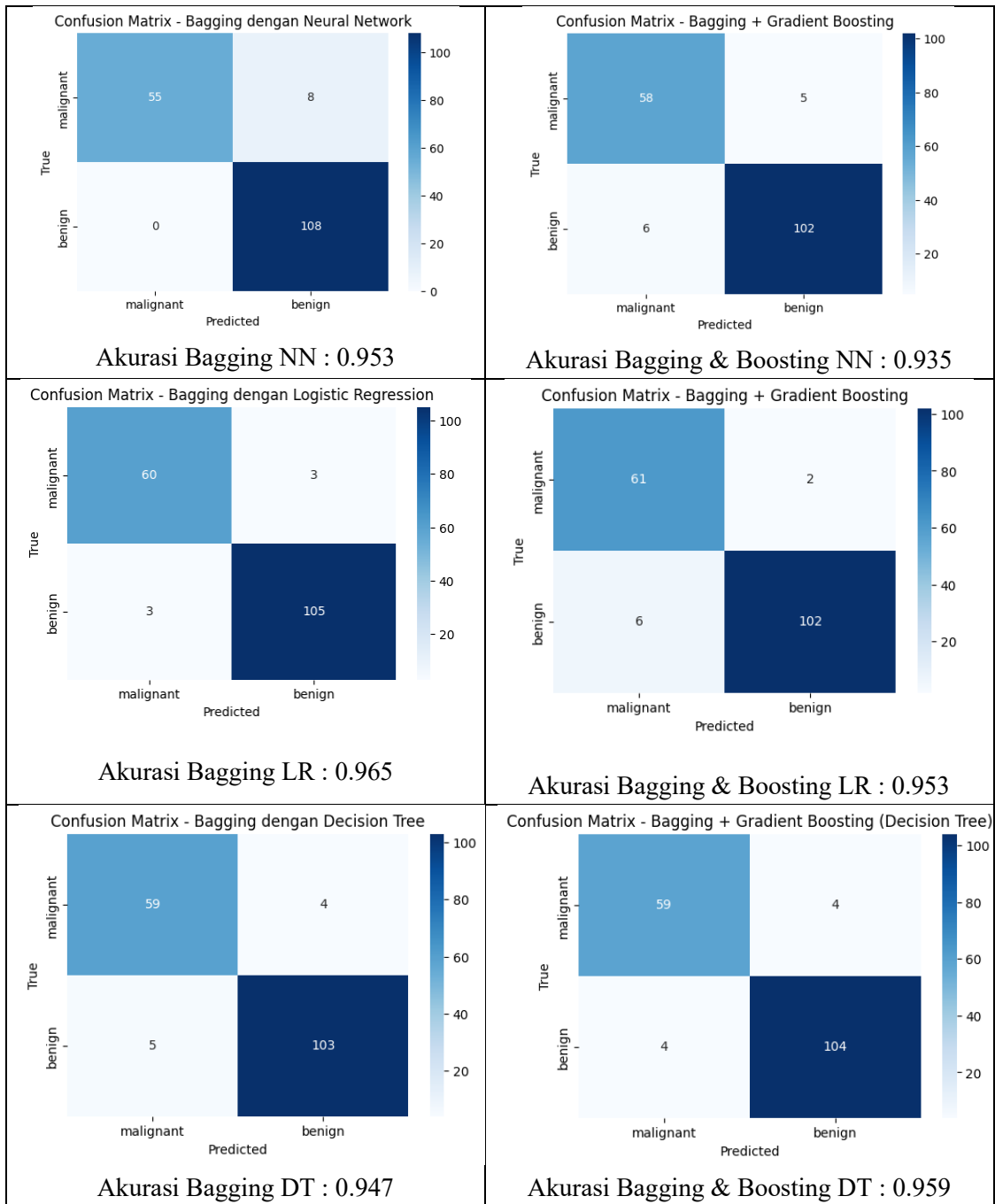
Tabel 1. Perbandingan Akurasi Model Bagging dan Gradient Boosting

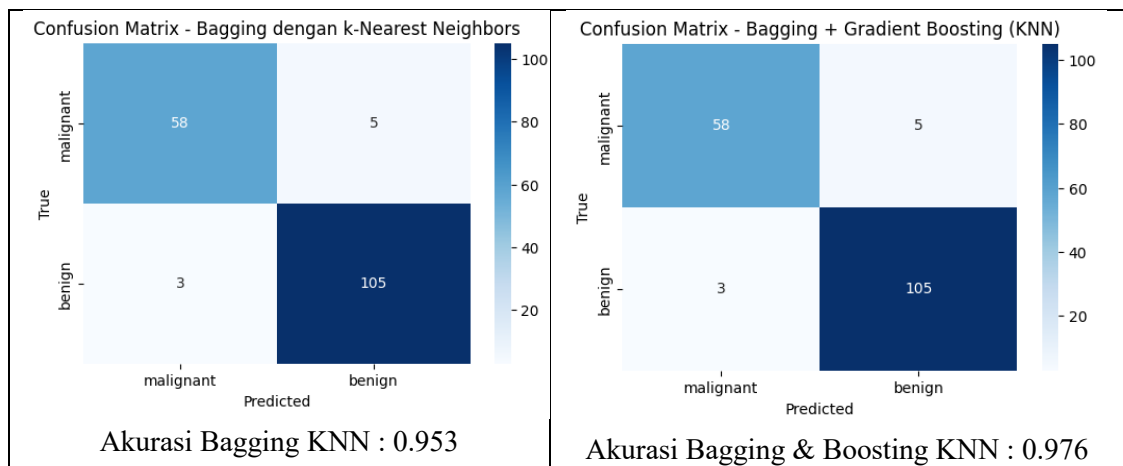
Model	Akurasi Bagging	Akurasi Bagging + Gradient Boosting
SVM	0.959	0.970
Neural Network	0.953	0.935
Logistic Regression	0.965	0.953
Decision Tree	0.947	0.959
K-Nearest Neighbors	0.953	0.976

Berdasarkan hasil pada Tabel 1 terlihat bahwa penerapan metode Gradient Boosting memberikan peningkatan performa pada beberapa algoritma klasifikasi. Peningkatan akurasi yang cukup signifikan terlihat pada algoritma K-Nearest Neighbors yang meningkat dari 0.953 menjadi 0.976. Selain itu, algoritma Support Vector Machine dan Decision Tree juga menunjukkan peningkatan performa setelah penerapan Gradient Boosting.

Sebaliknya, pada algoritma Neural Network dan Logistic Regression terjadi sedikit penurunan nilai akurasi setelah penerapan Gradient Boosting. Hal ini menunjukkan bahwa efektivitas metode ensemble dapat berbeda-beda tergantung pada karakteristik algoritma dasar yang digunakan.







Gambar 3. Perbandingan Akurasi Model Bagging dan Gradient Boosting

Gambar 3 menunjukkan visualisasi perbandingan nilai akurasi antara model yang menggunakan metode Bagging dan model yang menggunakan kombinasi Bagging dan Gradient Boosting pada lima algoritma klasifikasi.

#### b. Hasil AUC Model

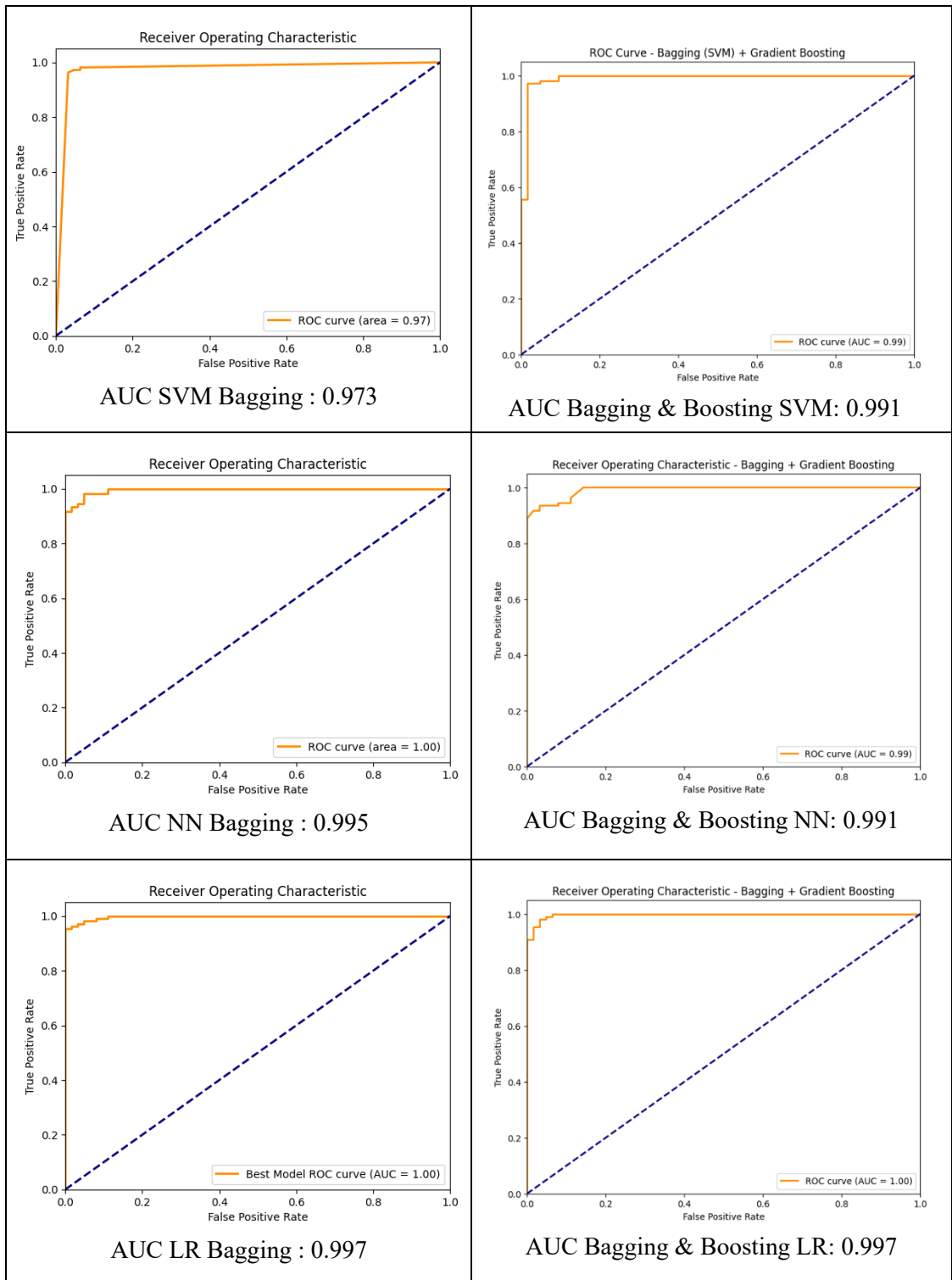
Selain akurasi, evaluasi performa model juga dilakukan menggunakan nilai Area Under Curve (AUC). Nilai AUC menunjukkan kemampuan model dalam membedakan antara kelas positif dan negatif.

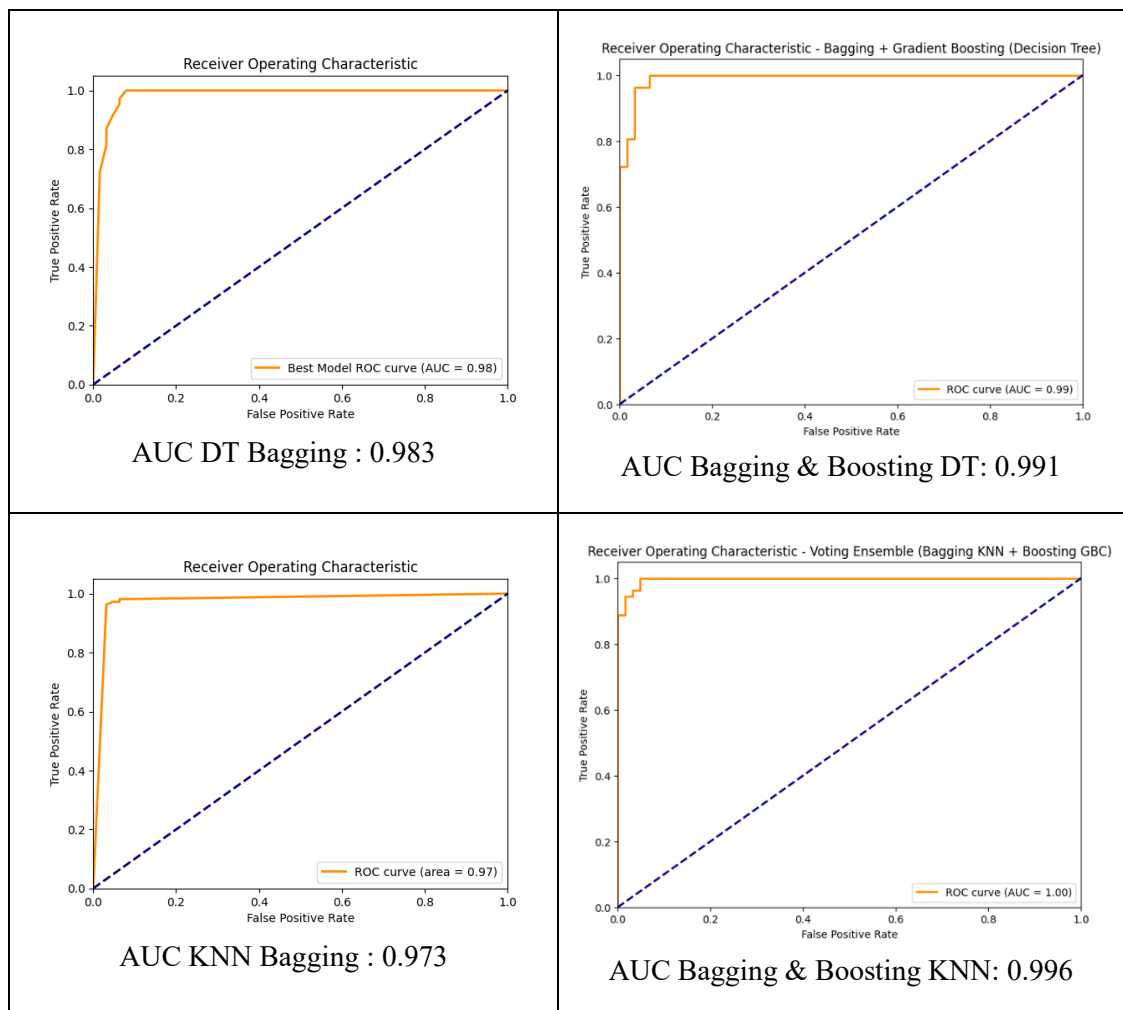
Tabel 2. Perbandingan Nilai AUC Model

Model	AUC Bagging	AUC Bagging + Gradient Boosting
SVM	0.973	0.991
Neural Network	0.995	0.991
Logistic Regression	0.997	0.997
Decision Tree	0.983	0.991
K-Nearest Neighbors	0.973	0.996

Berdasarkan Tabel 2 dapat dilihat bahwa sebagian besar algoritma mengalami peningkatan nilai AUC setelah penerapan metode Gradient Boosting. Nilai AUC yang tinggi menunjukkan bahwa model memiliki kemampuan yang baik dalam membedakan antara kelas positif dan negatif pada kasus diagnosis kanker payudara.

Algoritma K-Nearest Neighbors menunjukkan peningkatan AUC yang cukup signifikan dari 0.973 menjadi 0.996 setelah penerapan Gradient Boosting. Hal ini menunjukkan bahwa metode Boosting mampu meningkatkan kemampuan model dalam mempelajari pola data sehingga menghasilkan performa klasifikasi yang lebih baik.





Gambar 4. Kurva ROC Model Bagging dan Gradient Boosting

Gambar 4 menampilkan perbandingan kurva ROC antara model Bagging dan model Bagging yang dioptimasi menggunakan Gradient Boosting untuk setiap algoritma klasifikasi yang diuji.

## 2. PEMBAHASAN

Hasil eksperimen menunjukkan bahwa teknik ensemble learning memiliki pengaruh yang signifikan terhadap peningkatan performa model klasifikasi. Metode Bagging mampu meningkatkan stabilitas model dengan mengurangi variansi melalui pembentukan beberapa subset data bootstrap selama proses pelatihan.

Namun demikian, peningkatan performa yang lebih signifikan terlihat setelah diterapkannya metode Boosting. Pendekatan Gradient Boosting bekerja dengan membangun model secara bertahap untuk memperbaiki kesalahan prediksi yang terjadi pada model sebelumnya. Dengan mekanisme ini, model dapat mempelajari pola data secara lebih efektif sehingga menghasilkan performa klasifikasi yang lebih baik.

Berdasarkan hasil pengujian, Neural Network menunjukkan peningkatan performa yang paling signifikan setelah penerapan teknik ensemble learning. Hal ini menunjukkan bahwa model Neural Network memiliki kemampuan yang lebih baik

dalam memanfaatkan proses optimasi berbasis Boosting dibandingkan dengan beberapa algoritma lainnya.

Logistic Regression juga menunjukkan performa yang sangat stabil dengan nilai AUC sebesar 0.997 baik sebelum maupun sesudah penerapan Boosting. Hal ini menunjukkan bahwa algoritma Logistic Regression memiliki kemampuan klasifikasi yang baik pada dataset Breast Cancer Wisconsin.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa kombinasi metode Bagging dan Gradient Boosting mampu meningkatkan performa beberapa algoritma klasifikasi dalam kasus diagnosis kanker payudara. Teknik ensemble learning terbukti efektif dalam meningkatkan akurasi serta kemampuan model dalam membedakan kelas diagnosis sehingga berpotensi digunakan sebagai pendekatan yang lebih andal dalam pengembangan sistem diagnosis berbasis machine learning.

#### D. KESIMPULAN

Berdasarkan hasil evaluasi kinerja lima algoritma klasifikasi yaitu Support Vector Machine (SVM), Neural Network (NN), Logistic Regression (LR), Decision Tree (DT), dan K-Nearest Neighbors (KNN), penelitian ini menunjukkan bahwa penerapan pendekatan ensemble learning secara bertahap mampu meningkatkan performa prediksi pada beberapa model klasifikasi. Integrasi metode Bagging yang berfungsi untuk mengurangi variansi model, diikuti dengan penerapan Gradient Boosting yang berfokus pada perbaikan kesalahan prediksi, memberikan mekanisme optimasi yang lebih terstruktur dalam meningkatkan stabilitas serta kemampuan pemisahan kelas pada tugas klasifikasi medis.

Hasil penelitian menunjukkan bahwa peningkatan performa tidak hanya terjadi pada algoritma berbasis pohon keputusan, tetapi juga pada beberapa model klasifikasi lainnya. Hal ini menunjukkan bahwa penerapan Gradient Boosting dalam kerangka ensemble yang digunakan pada penelitian ini tidak terbatas pada jenis algoritma tertentu, melainkan dapat diterapkan pada berbagai model pembelajaran dengan karakteristik yang berbeda. Dengan demikian, pendekatan ensemble yang digunakan mampu memperluas pemanfaatan teknik boosting dalam berbagai arsitektur model klasifikasi.

Selain itu, peningkatan nilai akurasi dan Area Under Curve (AUC) pada beberapa model menunjukkan bahwa kombinasi metode Bagging dan Gradient Boosting tidak hanya meningkatkan ketepatan prediksi tetapi juga memperkuat kemampuan model dalam membedakan antara kelas malignant dan benign. Kemampuan pemisahan kelas yang baik menjadi faktor penting dalam konteks diagnosis medis karena berpengaruh langsung terhadap ketepatan pengambilan keputusan dalam proses diagnosis penyakit.

Secara keseluruhan, penelitian ini memberikan kontribusi dengan menunjukkan bahwa kombinasi teknik Bagging dan Gradient Boosting dapat digunakan sebagai strategi ensemble yang efektif untuk meningkatkan performa algoritma klasifikasi pada kasus diagnosis kanker payudara. Temuan ini diharapkan dapat menjadi referensi dalam pengembangan model klasifikasi berbasis machine learning pada bidang kesehatan serta membuka peluang penelitian lanjutan untuk mengeksplorasi penerapan metode ensemble pada berbagai dataset dan permasalahan klasifikasi lainnya.

#### E. DAFTAR PUSTAKA

- [1] H. J. Yoon *et al.*, "Accelerated training of bootstrap aggregation-based deep information extraction systems from cancer pathology reports," *J. Biomed. Inform.*, vol. 110, Oct. 2020, doi: 10.1016/j.jbi.2020.103564.

- 
- [2] A. H. Yunial, "Optimization Analysis of Support Vector Machine Classification Algorithms, Decision Trees, and Neural Networks Using Adaboost and Bagging," *J. Inform. Univ. Pamulang*, vol. 5, no. 3, p. 247, Sep. 2020, doi: 10.32493/informatika.v5i3.6609.
- [3] A. Kurniawan and A. Prihandono, "Application Of Bagging Techniques To Improve Classification Accuracy In The Naive Bayes Algorithm In Determining Professional Bloggers," 2020.
- [4] G. Arther Sandag, "A Prediction Model Of Company Health Using Bagging Classifier," vol. 6, no. 1, 2020, doi: 10.33480/jitk.v6i1.1390.
- [5] M. Dianati-nasab and K. Salimifard, "Machine learning algorithms to uncover risk factors of breast cancer : insights from a large case-control study," no. February, pp. 1–13, 2024, doi: 10.3389/fonc.2023.1276232.
- [6] H. Jafarzadeh, F. M. M. M. Eric Gill, and S. Homayouni, "Bagging and Boosting Ensemble Classifiers for Classification of Comparative Evaluation," *Remote sensing, MDPI*, 2021.
- [7] M. Dutta, K. M. Hasan, A. Akter, and H. Rahman, "An interpretable machine learning-based breast cancer classification using XGBoost , SHAP , and LIME," vol. 13, no. 6, pp. 4306–4315, 2024, doi: 10.11591/eei.v13i6.7866.
- [8] R. du Plooy and P. J. Venter, "A Comparison of Artificial Neural Networks and Bootstrap Aggregating Ensembles in a Modern Financial Derivative Pricing Framework," *J. Risk Financ. Manag.*, vol. 14, no. 6, p. 254, Jun. 2021, doi: 10.3390/jrfm14060254.
- [9] V. I. Yani, A. Aradea, and H. Mubarak, "Optimasi Prakiraan Cuaca Menggunakan Metode Ensemble pada Naïve Bayes dan C4.5," *J. Tek. Inform. dan Sist. Inf.*, vol. 8, no. 3, pp. 607–619, 2022, doi: 10.28932/jutisi.v8i3.5455.
- [10] L. M. Cendani and A. Wibowo, "Perbandingan Metode Ensemble Learning pada Klasifikasi Penyakit Diabetes," *J. Masy. Inform.*, vol. 13, no. 1, pp. 33–44, 2022, doi: 10.14710/jmasif.13.1.42912.
- [11] R. Pramudita, S. Muis, N. Safitri, and F. Shafirawati, "Optimasi Algoritma Machine Learning Menggunakan Teknik Bagging Pada Klasifikasi Diagnosis Kanker Payudara," *Tematik*, vol. 11, no. 1, pp. 128–134, 2024, doi: 10.38204/tematik.v11i1.1928.
- [12] S. Al-Dahidi, P. Baraldi, E. Zio, and M. Lorenzo, "Bootstrapped ensemble of artificial neural networks technique for quantifying uncertainty in prediction of wind energy production," *Sustain.*, vol. 13, no. 11, Jun. 2021, doi: 10.3390/su13116417.
- [13] C. O. Siti Fatimah *et al.*, "The effect of aggregating bootstrap on the accuracy of neural network system for islamic investment prediction," *Univers. J. Account. Financ.*, vol. 9, no. 4, pp. 604–612, Aug. 2021, doi: 10.13189/ujaf.2021.090408.
- [14] R. Pramudita, Solikin, and N. Safitri, "Optimization Analysis of Neural Network Algorithms Using Bagging Techniques on Classification of Date Fruit Types," in *Seventh International Conference on Informatics and Computing (ICIC)*, Denpasar: IEEE, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/10006986>
- [15] R. Ali and A. Shabri, "Comparative analysis of hybrid models for prediction of tourist arrivals," in *AIP Conference Proceedings*, S. A., E. A.F., H. S.P., R. H.A., A. M.A.K., and S. S., Eds., American Institute of Physics Inc., 2023. doi: 10.1063/5.0114935.