

Terbit online pada laman web jurnal: <https://jurnal.plb.ac.id/index.php/tematik/index>

T E M A T I K

Jurnal Teknologi Informasi Komunikasi (e-Journal)

Vol. 11 No. 1 (2024) 128 - 134

ISSN Media Elektronik: 2443-3640

Optimasi Algoritma Machine Learning Menggunakan Teknik Bagging Pada Klasifikasi Diagnosis Kanker Payudara

Optimization of Machine Learning Algorithms Using Bagging Techniques in Breast Cancer Diagnosis Classification

Rully Pramudita^{1*}, Saludin Muis², Nadya Safitri³, Fitri Shafirawati⁴^{1,2,4}Teknik Informatika, Fakultas Informatika, Universitas Bina Insani³Rekayasa Prangkat Lunak, Fakultas Informatika, Universitas Bina Insani¹rullypramudita@binainsani.ac.id, ²saludin@binainsani.ac.id, ³nadyasafitri@binainsani.ac.id, ⁴fshafira0904@gmail.com**Abstract**

Classification algorithms have a very important role in Machine Learning, but not all algorithms have the same performance in every case. Algorithm performance can be affected by the type of data used, differences in problem characteristics, and the parameters used. Additionally, ensemble learning techniques such as Bagging can affect algorithm performance. Therefore, the problem arises of how to choose the most suitable algorithm for a particular classification task and how to optimize the performance of the algorithm. This research aims to carry out a comparative analysis and optimization of classification algorithms in Machine Learning. Classification algorithms that will be evaluated include Support Vector Machine (SVM), Neural Network, Logistic Regression, Decision Tree, and K-Nearest Neighbors (K-NN). Evaluation of the performance of these algorithms will be carried out using the confusion matrix, Receiver Operating Characteristic (ROC) Curve, and Area Under Curva (AUC). The result of this research is a comparative analysis of the optimization of classification algorithms using the bagging technique. After carrying out the evaluation process using the confusion matrix and ROC curve, it was found that the algorithm optimization using the bagging technique only had an effect on the Decision Tree (DT) and K-Nearest Neighbors (KNN) algorithms. The accuracy of the DT algorithm increased by 0.6% while the accuracy of KNN increased by 1.3%. The AUC value for the DT algorithm increased by 1.4% and the KNN algorithm increased by 0.3%.

Keywords: algorithms, bagging, classification, optimization

Abstrak

Algoritma klasifikasi memiliki peran yang sangat penting dalam Machine Learning, namun tidak semua algoritma memiliki kinerja yang sama dalam setiap kasus. Kinerja algoritma dapat dipengaruhi oleh jenis data yang digunakan, perbedaan dalam karakteristik masalah, dan parameter yang digunakan. Selain itu, teknik ensemble learning seperti Bagging dapat memengaruhi kinerja algoritma. Oleh karena itu, muncul permasalahan tentang bagaimana memilih algoritma yang paling sesuai untuk tugas klasifikasi tertentu dan bagaimana mengoptimalkan kinerja algoritma tersebut. Penelitian ini bertujuan untuk melakukan analisis perbandingan dan optimasi algoritma klasifikasi pada Machine Learning. Algoritma klasifikasi yang akan dievaluasi termasuk Support Vector Machine (SVM), Neural Network, Logistic Regression, Decision Tree, dan K-Nearest Neighbors (K-NN). Evaluasi kinerja algoritma-algoritma ini akan dilakukan menggunakan confusion matrix, Kurva Receiver Operating Characteristic (ROC), dan Area Under Curva (AUC). Hasil dari penelitian ini adalah sebuah analisis perbandingan optimasi algoritma klasifikasi menggunakan teknik bagging, setelah dilakukan proses evaluasi menggunakan confusion matriks dan kurva ROC, dihasilkan bahwa optimasi algoritma menggunakan teknik bagging hanya berpengaruh pada algoritma Decision Tree (DT) dan K-Nearest Neighbors (KNN). Akurasi algoritma DT meningkat sebesar 0.6% sedangkan akurasi KNN meningkat sebesar 1.3%. Untuk nilai AUC algoritma DT meningkat sebesar 1.4% dan algoritma KNN meningkat sebesar 0.3%.

Kata kunci: algoritma, bagging, klasifikasi, optimasi

1. Pendahuluan

Machine Learning telah merevolusi cara memahami dan mengambil keputusan berdasarkan data. Salah satu

tugas penting dalam Machine Learning adalah klasifikasi, di mana kita memprediksi kategori atau label dari suatu data berdasarkan fitur-fitur yang ada. Kemampuan ini sangat penting dalam berbagai aplikasi,

termasuk pengenalan pola, klasifikasi medis, rekomendasi produk, dan masih banyak lagi. Algoritma klasifikasi adalah inti dari Machine Learning, dan berbagai jenis algoritma telah dikembangkan untuk menangani tugas ini.

Meskipun algoritma klasifikasi memiliki peran yang sangat penting dalam Machine Learning, tidak semua algoritma memiliki kinerja yang sama dalam setiap kasus. Kinerja algoritma dapat dipengaruhi oleh jenis data yang digunakan, perbedaan dalam karakteristik masalah, dan parameter yang digunakan. Selain itu, teknik ensemble learning seperti Bagging dapat memengaruhi kinerja algoritma. Oleh karena itu, muncul permasalahan tentang bagaimana memilih algoritma yang paling sesuai untuk tugas klasifikasi tertentu dan bagaimana mengoptimalkan kinerja algoritma tersebut.

Pada penelitian sebelumnya Green Arther [1] melakukan prediksi company health menggunakan perbandingan bagging dan Xgboost menggunakan model Algoritma Dasar SVM, Logistic Regression, dan Decision Tree dengan hasil akhir bagging memiliki hasil akurasi yang lebih baik. Penelitian lainnya dari Hong Jun Yoon et al membuktikan bahwa agregat model meningkatkan performansi jika dibandingkan dengan penggunaan model tunggal. Dan menunjukkan bahwa metode bagging memiliki akurasi klasifikasi yang lebih tinggi dari model dasarnya, dimana model dasar yang digunakan yaitu MT-CNN dan MT-HCAN [2]. Penelitian ketiga yang telah dilakukan oleh [3], melakukan analisis optimasi menggunakan adaboost dan bagging dari model dasar SVM, decision tree dan NN, Dalam penelitian ini, ditemukan bahwa dengan mengoptimalkan Adaboost, terjadi peningkatan signifikan pada akurasi beberapa algoritma. Akurasi algoritma Support Vector Machine meningkat dari 88.93% menjadi 89.10%, Decision Trees dari 90.24% menjadi 90.36%, dan Neural Network dari 88.53% menjadi 88.61%. Sementara itu, untuk algoritma Bagging, hanya terjadi peningkatan pada Decision Trees, mencapai 90.55%, dan Neural Network, mencapai 90.38%. Nilai akurasi algoritma Support Vector Machine tetap sama dengan nilai akurasi Bagging, yaitu 88.93%. Keempat penelitian dari [4][5] menggunakan bagging dan model dasar algoritma Naïve Bayes, berdasarkan temuan penelitiannya, penerapan teknik bagging dalam klasifikasi ensemble pada algoritma Naïve Bayes menghasilkan peningkatan akurasi. Kelima, penelitian dari [6] menemukan bahwa jaringan ensemble agregasi bootstrap mengungguli jaringan saraf tiruan dan menghasilkan perkiraan harga dalam batas kesalahan Monte Carlo simulasi saat menetapkan harga derivatif dalam pengaturan kerangka kerja multi-kurva. Keenam, pada penelitian [7] telah menghasilkan hasil yang menunjukkan bahwa teknik BS (Bootstrap) memberikan penilaian yang lebih memuaskan terhadap tingkat ketidakpastian dalam

prediksi energi angin daripada teknik yang digunakan oleh pemilik pembangkit angin tersebut dan juga lebih baik daripada teknik Estimasi Rerata-Variansi (MVE) yang ada dalam literatur. Hasil dari teknik BS juga dianalisis dalam konteks kondisi cuaca yang berbeda yang dialami oleh pembangkit angin dan dalam jangka waktu prediksi yang beragam. Ketujuh [8] menyatakan bahwa penggunaan metode Aggregating Bootstrap terbukti efektif dalam mengurangi bias dalam sistem jaringan saraf (neural network), yang pada akhirnya meningkatkan akurasi prediksi. Kedelapan pada penelitian [9][10] menyebutkan hasil evaluasi model menunjukkan bahwa penggunaan algoritma neural network bersama teknik bagging menghasilkan tingkat akurasi lebih baik dibandingkan dengan penggunaan algoritma neural network secara tunggal. Kesembilan hasil eksperimen [11] menunjukkan bahwa pengoptimalan dengan menggunakan metode ensemble Bagging pada algoritma Naïve Bayes mengakibatkan peningkatan akurasi, Algoritma C4.5, di sisi lain, mengalami peningkatan akurasi yang lebih signifikan setelah dioptimalkan menggunakan metode ensemble. Pengoptimalan dengan metode ensemble Bagging pada algoritma C4.5 menghasilkan peningkatan akurasi [12]. Pada penelitian [13] dilakukan perbandingan ensemble learning pada 4 algoritma dasar dan menghasilkan kesimpulan bahwa bagging metode terbaik untuk ensemble learning. Perbandingan metode ensemble learning yang dilakukan oleh [14] menyatakan bahwa metode Boosting mengungguli metode Bagging dan Stacking. Pada dataset pertama, dicapai tingkat akurasi tertinggi sebesar 81,82% dengan menggunakan model-model seperti Gradient Boosting, Extreme Gradient Boosting, dan Cat Boosting. Pada dataset kedua, tingkat akurasi tertinggi mencapai 99,25% dengan pemanfaatan model Light Gradient Boosting. Sementara itu, pada dataset ketiga, terdapat pencapaian akurasi maksimal sebesar 100% dengan memanfaatkan model-model Light Gradient Boosting dan Cat Boosting. Algoritma Bagging memiliki dampak yang lebih stabil pada model yang digunakan pada penelitian prediksi Dst Index [15]. Pada penelitian [16][17][18][19][20] menyatakan bahwa pada penelitian tersebut bagging atau bootstraps memiliki dampak positif dalam meningkatkan performa model yang diusulkan dalam masing-masing kasusnya. Berdasarkan penelitian-penelitian sebelumnya dapat disimpulkan bahwa dari penelitian sebelumnya menggunakan teknik ensemble dalam melakukan optimasi algoritma. Dalam hal model tunggal yang digunakan sangat beragam diantaranya SVM, Logistic Regression, Decision Tree, Neural Network, ANN, Naïve Bayes, MT-CNN dan MT-HCAN. Namun rata-rata paling banyak hanya menggunakan 3 sampai 4 model tunggal. Maka gap permasalahan yang muncul yaitu pertama belum adanya perbandingan lebih dari 4 model tunggal, kedua belum ada yang membandingkan algoritma Support Vector Machine (SVM), Neural

Network, Logistic Regression, Decision Tree, dan K-Nearest Neighbors (K-NN) pada penelitian sebelumnya. Maka penelitian ini akan dilakukan sebuah percobaan menggunakan 5 algoritma sebagai model tunggal. Selain itu pada penelitian-penelitian sebelumnya mengatakan bahwa Bootstraps Aggregating (Bagging) dapat meningkatkan performance dari model tunggal. Berdasarkan beberapa hal tersebut, maka akan menjadi dasar tujuan penelitian ini melakukan percobaan analisis perbandingan algoritma klasifikasi menggunakan 5 model tunggal Support Vector Machine (SVM), Neural Network, Logistic Regression, Decision Tree, dan K-Nearest Neighbors (K-NN) serta menggunakan algoritma bagging sebagai algoritma optimasi atau ensemble. Dan dievaluasi menggunakan confusion metrics, ROC dan AUC.

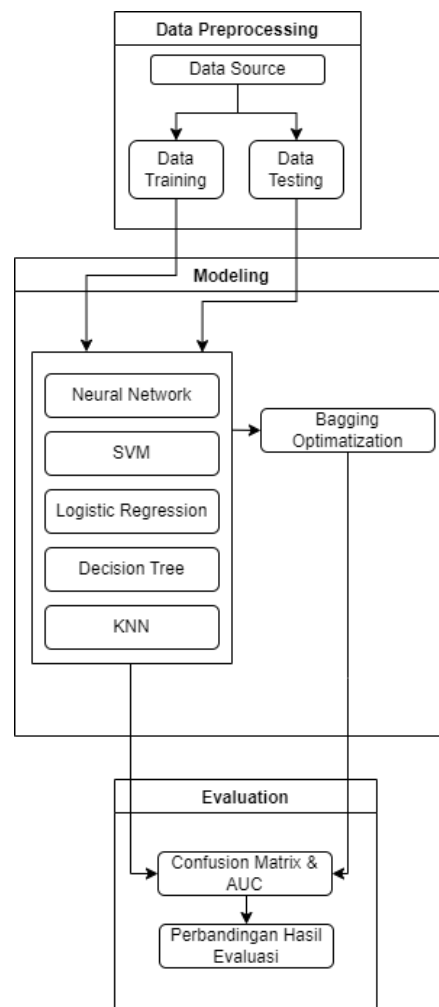
2. Metode Penelitian

Penelitian ini menggunakan dataset publik yang dikenal sebagai "Breast Cancer Wisconsin (Diagnostic)" yang berasal dari platform Kaggle. Dataset ini telah menjadi salah satu sumber data yang penting dalam penelitian medis dan Machine Learning, terutama dalam konteks klasifikasi kanker payudara. Data ini mencakup berbagai fitur medis yang berkaitan dengan jaringan payudara dan diagnosis kanker payudara. Dengan menggunakan dataset ini, penelitian memiliki akses ke informasi yang penting untuk analisis klasifikasi yang akurat dan membantu dalam pengembangan model-machine learning yang efektif dalam mendukung diagnosis dini kanker payudara. Proses tahapan penelitian ini dilakukan berdasarkan Gambar 1.

Berdasarkan Gambar 1, penelitian dimulai dari tahapan data preprocessing, pada tahapan ini dilakukan proses pengumpulan data source yang akan digunakan dalam penelitian. Karakteristik dari dataset tersebut ialah berjumlah 569 record, dengan 30 atribut numerik, label dari dataset ini terdapat 2 label yaitu "malignant" dan "benign" tanpa missing attribute values. Sehingga dataset sudah dapat dikatakan siap diproses ketahap pemodelan. Setelah data source didapatkan maka langkah selanjutnya adalah melakukan pembagian data menjadi data training dan data test dengan perbandingan 70:30. Hasil dari proses ini akan menjadi masukan untuk tahapan modeling data.

Tahapan kedua yaitu dilakukan modeling data, pada proses ini akan dilakukan pemodelan data berdasarkan data training dan data testing yang dihasilkan pada tahapan pertama. Pemodelan tahap 1 akan dilakukan pendefinisian model dasar yang akan digunakan Support Vector Machine (SVM), Neural Network, Logistic Regression, Decision Tree, dan K-Nearest Neighbors (K-NN) dan melakukan pelatihan data pada model dasar. Pemodelan tahap 2 akan dilakukan pendefinisian dan pelatihan model bagging dengan

model dasar secara bergantian dan diakhiri dengan proses aggregating.



Gambar 1. Tahapan Penelitian

Tahap terakhir pada penelitian ini akan dilakukan proses evaluation, proses evaluasi dilakukan untuk menghitung confusion metrics, ROC dan AUC dari masing-masing model dasar dan hasil penggunaan algoritma bagging. Jika semua nilai sudah berhasil dihitung maka langkah akhir yaitu melakukan analisis perbandingan nilai akurasi dan AUC.

3. Hasil dan Pembahasan

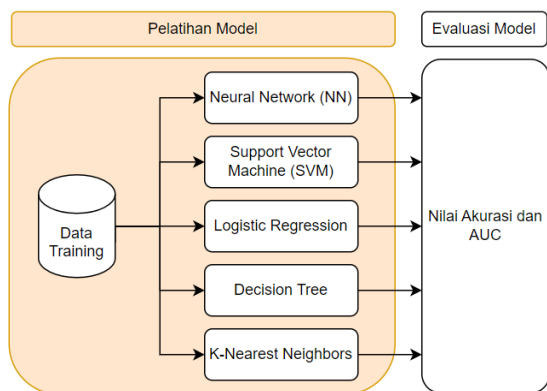
Karakteristik dari dataset tersebut ialah berjumlah 569 record, dengan 30 atribut numerik, label dari dataset ini terdapat 2 label yaitu "malignant" dan "benign" tanpa missing attribute values. Sehingga dataset sudah dapat dikatakan siap diproses ketahap pemodelan. Setelah data source disiapkan maka langkah selanjutnya adalah melakukan pembagian data menjadi data training dan data test dengan perbandingan 70:30. Hasil dari proses ini akan menjadi masukan untuk tahapan modeling data.

Rangkaian hasil penelitian berdasarkan urutan/susunan logis untuk membentuk sebuah cerite. Isinya menunjukkan fakta/data dan jangan diskusikan hasilnya. Dapat menggunakan Tabel dan Angka tetapi tidak menguraikan secara berulang terhadap data yang sama dalam gambar, tabel dan teks. Untuk lebih memperjelas uraian, dapat menggunakan sub judul.

Pembahasan adalah penjelasan dasar, hubungan dan generalisasi yang ditunjukkan oleh hasil. Uraianya menjawab pertanyaan penelitian. Jika ada hasil yang meragukan maka tampilkan secara objektif.

Pada tahapan Data Preprocessing ini dilakukan proses pengumpulan data source yang akan digunakan dalam penelitian. Dataset yang digunakan yaitu Breast cancer wisconsin (diagnostic) dari data public pada kaggle dataset. Karakteristik dari dataset tersebut ialah berjumlah 569 record, dengan 30 atribut numerik, label dari dataset ini terdapat 2 label yaitu “malignant” dan “benign” tanpa missing attribute values. Sehingga dataset sudah dapat dikatakan siap diproses ketahap pemodelan. Setelah data source disiapkan maka langkah selanjutnya adalah melakukan pembagian data menjadi data training dan data test dengan perbandingan 70:30. Hasil dari proses ini akan menjadi masukan untuk tahapan modeling data.

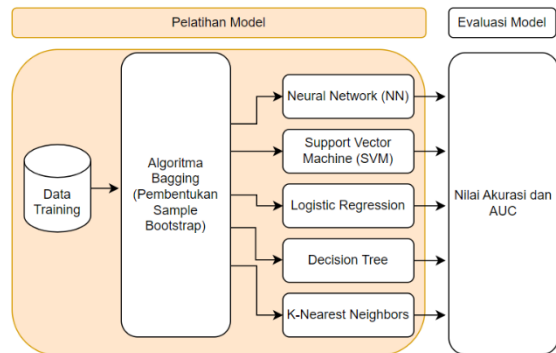
Pemodelan Tahap 1 dilakukan untuk pelatihan model algoritma dasar klasifikasi menggunakan data training. Pelatihan dilakukan secara satu persatu dari model dasar klasifikasi mulai dari Support Vector Machine (SVM), Neural Network (NN), Logistic Regression (LR), Decision Tree (DT), dan K-Nearest Neighbors (K-NN). Secara teknis proses pelatihan model-model dasar tersebut dapat dilihat pada kode program pyhton pada halaman lampiran. Gambar 2 menggambarkan bagaimana data training digunakan untuk inputan melatih setiap modelnya dan di akhir akan diuji pada tahap evaluasi.



Gambar 2. Pelatihan Model Algoritma Dasar

Setelah melakukan pemodelan 1 untuk melatih setiap model algoritma dasar, maka pada pemodelan 2 ini akan mulai dilakukan penggunaan algoritma bagging dalam melatih setiap model dasarnya. Pertama-tama data

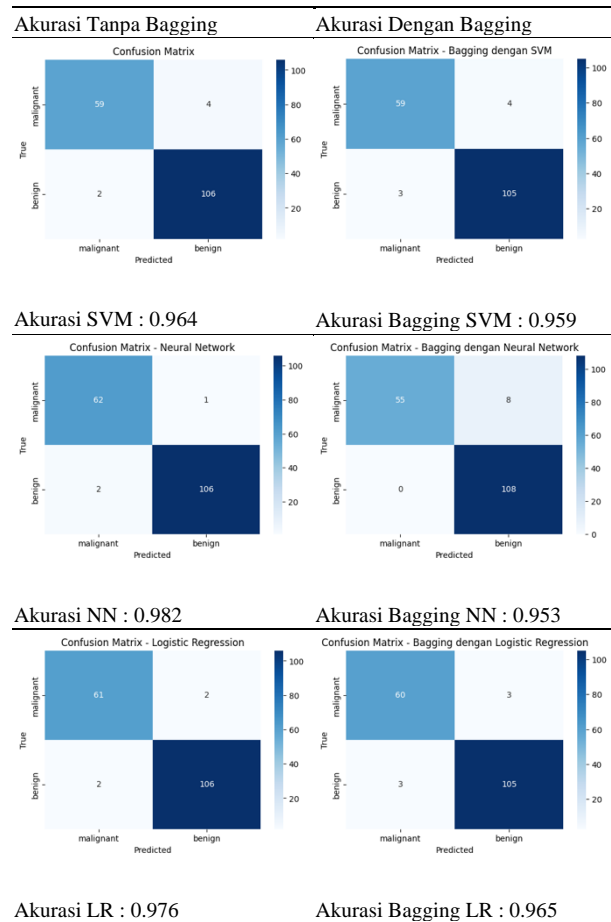
training akan dibagi menjadi beberapa sample bootstrap oleh algoritma bagging, yang selanjutnya sampling tersebut akan digunakan oleh setiap model dasar. Gambar 3 menggambarkan bagaimana penggunaan algoritma bagging disaat melakukan pelatihan model.

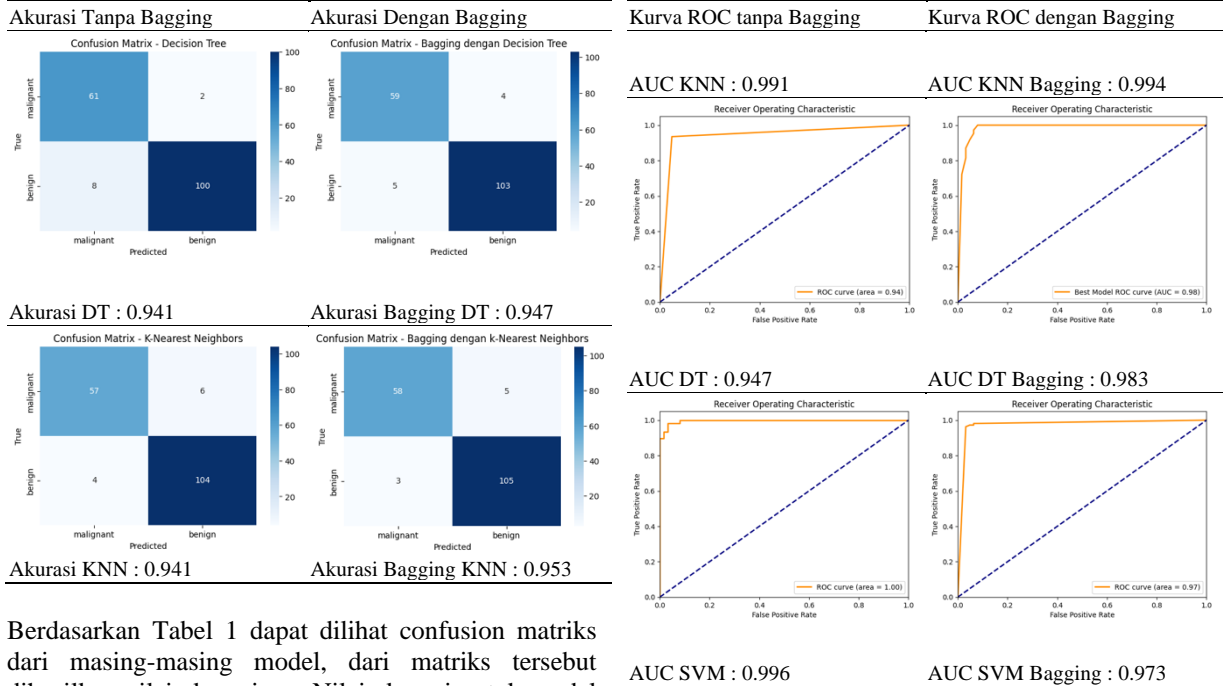


Gambar 3 Pelatihan Model Algoritma Dasar Ditambah Algoritma Bagging

Berdasarkan proses evaluasi setiap model menggunakan confusion matriks dan Kurva ROC, maka dihasilkan sebuah perbandingan antar model yang dilihat dari nilai akurasi dan nilai AUC nya dari masing-masing model. Pada Tabel 1 tergambaran confusion matriks dan nilai akurasi dari setiap modelnya.

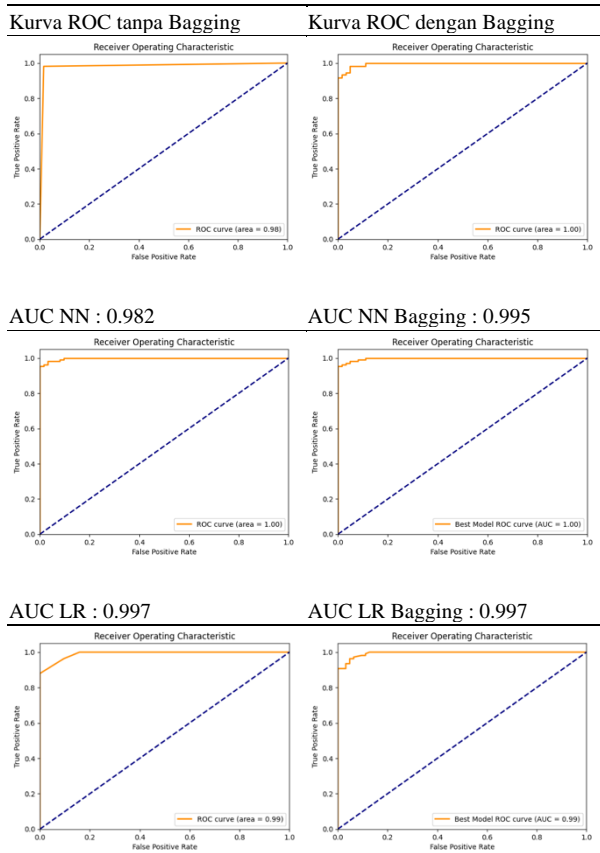
Tabel 1. Perbandingan Akurasi





Berdasarkan Tabel 1 dapat dilihat confusion matriks dari masing-masing model, dari matriks tersebut dihasilkan nilai akurasi. Nilai akurasi untuk model dasar yang paling tinggi dihasilkan dari model Neural Network (NN), sedangkan nilai akurasi setelah dilakukan bagging, yang tertinggi dihasilkan dari model Logistic Regression (LR).

Tabel 2 Perbandingan Kurva ROC



Selain menggunakan confusion matriks, evaluasi pun dilakukan menggunakan kurva ROC untuk melihat performance dari masing-masing model yang diujikan. Pada Tabel 2 dapat dilihat kurva ROC dari masing-masing model. Kurva tersebut menghasilkan nilai AUC yang menunjukkan performance dari masing-masing model. Nilai AUC paling tinggi pada model dasar yaitu Logistic Regression (LR). Sedangkan nilai AUC paling tinggi pada model yang telah ditambahkan bagging yaitu:

Tabel 3. Hasil Perhitungan Akurasi dan AUC

Model	Akurasi		Ket	AUC		Ket
	Tanpa Bagging	Dengan Bagging		Tanpa Bagging	Dengan Bagging	
SVM	96,4%	95,9%	↓	99,6%	97,3%	↓
NN	98,2%	95,3%	↓	98,2%	99,5%	↑
LR	97,6%	96,5%	↓	99,7%	99,7%	-
DT	94,1%	94,7%	↑	94,7%	98,3%	↑
KNN	94,1%	95,3%	↑	99,1%	99,4%	↑

Berdasarkan Tabel 3 hasil perbandingan dari perhitungan akurasi dan AUC, untuk nilai akurasi dan AUC yang terjadi peningkatan dengan penggunaan bagging yaitu pada algoritma Decision Tree (DT) dan K-Nearest Neighbors (KNN). Untuk Neural Network (NN) hanya mengalami peningkatan dari segi performance yang terlihat dari nilai AUC. Sedangkan algoritma lain tidak terjadi peningkatan dengan penggunaan bagging. Maka dapat disimpulkan penemuan dari penelitian ini membuktikan bahwa algoritma bagging dapat diterapkan secara optimal dalam hal meningkatkan akurasi serta performance hanya pada algoritma Decision Tree dan K-Nearest Neighbors serta peningkatan nilai performance pada algoritma Neural Network (NN). Hasil ini ada

kesamaan dalam peningkatan nilai akurasi pada Decision Tree menggunakan teknik bagging dengan penelitian yang dilakukan oleh [3] dan menambahkan peningkatan nilai performance pada algoritma Neural Network.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan dimulai dari tahapan data preprocessing hingga evaluasi. Berikut kesimpulan yang didapatkan, penelitian ini berhasil melakukan klasifikasi menggunakan algoritma-klasifikasi SVM, Neural Network, Logistic Regression, Decision Tree, dan K-NN dengan menggunakan google colab, penelitian ini berhasil melakukan analisis kinerja masing-masing algoritma klasifikasi (SVM, Neural Network, Logistic Regression, Decision Tree, dan K-NN) dalam hal akurasi dan performance algoritma menggunakan confusion matriks dan kurva ROC, berhasil melakukan analisis pengaruh optimasi terhadap akurasi dan performance algoritma, dan ternyata optimasi menggunakan teknik bagging tidak berpengaruh terhadap semua algoritma klasifikasi. Optimasi menggunakan teknik bagging hanya berpengaruh pada algoritma Decision Tree (DT) dan K-Nearest Neighbors (KNN). Akurasi algoritma DT meningkat sebesar 0.6% sedangkan akurasi KNN meningkat sebesar 1.3%. Untuk nilai AUC algoritma DT meningkat sebesar 1.4% dan algoritma KNN meningkat sebesar 0.3%, berhasil melakukan pengukuran tingkat akurasi model algoritma klasifikasi SVM, Neural Network, Logistic Regression, Decision Tree, dan K-NN menggunakan Confusion Matriks, berhasil melakukan pengukuran performance model algoritma klasifikasi SVM, Neural Network, Logistic Regression, Decision Tree, dan K-NN menggunakan AUC. Adapun saran untuk penelitian selanjutnya, penggunaan teknik ensemble learning yang lain untuk melakukan optimasi algoritma, seperti Boosting atau Stacking, penggunaan dataset lain yang memiliki karakter dan jumlah data yang berbeda, penggunaan algoritma klasifikasi lain untuk melakukan analisis perbandingan algoritma, penggunaan teknik evaluasi lain untuk mengevaluasi model algoritma.

Ucapan Terimakasih

Lembaga Penelitian dan Pengabdian Masyarakat Universitas Bina Insani.

Daftar Rujukan

- [1] G. Arther Sandag, "A Prediction Model Of Company Health Using Bagging Classifier," vol. 6, no. 1, 2020, doi: 10.33480/jitk.v6i1.1390.
- [2] 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.
- [3] 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.
- [4] A. Kurniawan and A. Prihandono, "Application Of Bagging Techniques To Improve Classification Accuracy In The Naive Bayes Algorithm In Determining Professional Bloggers," 2020.
- [5] Y. L. Nursimpati and A. Saifudin, "Application of Naïve Bayes-Based Bagging Technique for Student Admissions Selection," vol. 4, no. 2, pp. 2622–4615, 2019, [Online]. Available: <http://openjournal.unpam.ac.id/index.php/informatika65>
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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>
- [10] J. Abdollahi, B. Nouri-Moghaddam, and M. Ghazanfari, "Deep Neural Network Based Ensemble learning Algorithms for the healthcare system (diagnosis of chronic diseases)," 2021. doi: <https://doi.org/10.48550/arXiv.2103.08182>.
- [11] 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.
- [12] S. Mutmainnah, G. Abdurrahman, H. Azizah, and A. Faruq, "C4.5 Algorithm Optimization Using Bagging Techniques On Gold Karat Level Data," 2020. Accessed: Nov. 29, 2022. [Online]. Available: <http://repository.unmuhjember.ac.id/4416/>
- [13] R. Singh, "Machine Learning Algorithms and Ensemble Technique to Improve Prediction of Students Performance," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 3, pp. 3970–3976, Jun. 2020, doi: 10.30534/ijatcse/2020/221932020.
- [14] 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.
- [15] S. B. Xu, S. Y. Huang, Z. G. Yuan, X. H. Deng, and K. Jiang, "Prediction of the Dst Index with Bagging Ensemble-learning Algorithm," *Astrophys. J. Suppl. Ser.*, vol. 248, no. 1, p. 14, May 2020, doi: 10.3847/1538-4365/ab880e.
- [16] I. Possebon, A. Silva, L. Granville, A. Schaeffer-Filho, and A. Marnerides, "Improved Network Traffic Classification Using Ensemble Learning," *IEEE Symp. Comput. Commun.*, 2019.
- [17] D. Triana and S. Osowski, "Bagging and boosting techniques in prediction of particulate matters," *Bull. Polish Acad. Sci. Tech. Sci.*, vol. 68, no. 4, pp. 1207–1215, Oct. 2020, doi: 10.24425/bpasts.2020.134659.
- [18] A. Dželihodžić, D. Donko, and J. Kevrić, "Improved credit scoring model based on bagging neural network," *Int. J. Inf. Technol. Decis. Mak.*, vol. 17, no. 6, pp. 1725–1741, Nov. 2018, doi: 10.1142/S0219622018500293.
- [19] M. Smaida, "Bagging of Convolutional Neural Networks for Diagnostic of Eye Diseases," 2020.
- [20] U. Michelucci and F. Venturini, "Estimating Neural Network's

Performance with Bootstrap: A Tutorial,” *Mach. Learn. Knowl. Extr.*, vol. 3, no. 2, pp. 357–373, Mar. 2021, doi:

10.3390/make3020018.