

Klasifikasi Diagnosis Penyakit Alzheimer Menggunakan Adasyn Dan Algoritma Multilayer Perceptron Classifier

Fauzan Azim^{1*}, Yuli Hartati², Ipriadi³, Rahmad Gunawan¹

¹ Program Studi Teknik Informatika, Universitas Muhammadiyah Riau, Indonesia.

² Program Studi Sistem Informasi, Institut Teknologi dan Ilmu Sosial Khatulistiwa, Indonesia.

³ Program Studi Sistem Komputer, Universitas Putra Indonesia "YPTK" Padang, Indonesia.

Artikel Info

Kata Kunci:

ADASYN;
Alzheimer;
Klasifikasi Diagnosis.

Keywords:

ADASYN;
Alzheimer;
Diagnostic Classification.

Riwayat Artikel:

Submitted: 11 Juni 2026

Accepted: 28 September 2026

Published: 30 September 2026

Abstrak: Penyakit Alzheimer merupakan salah satu gangguan neurodegeneratif yang menjadi penyebab utama demensia dan ditandai dengan penurunan fungsi kognitif secara progresif. Seiring meningkatnya jumlah penderita Alzheimer, diperlukan metode klasifikasi yang mampu mendukung proses diagnosis secara cepat dan akurat. Salah satu permasalahan yang sering ditemukan pada data medis adalah ketidakseimbangan distribusi kelas yang dapat menurunkan kinerja model klasifikasi. Penelitian ini bertujuan untuk meningkatkan performa klasifikasi diagnosis penyakit Alzheimer dengan menerapkan metode Adaptive Synthetic Sampling (ADASYN) dan algoritma Multilayer Perceptron (MLP). Dataset yang digunakan diperoleh dari Kaggle dan terdiri atas data demografi, kondisi kesehatan, aktivitas harian, serta indikator klinis pasien. Tahapan penelitian meliputi preprocessing data, seleksi fitur menggunakan Information Gain, penyeimbangan data menggunakan ADASYN, pemodelan menggunakan MLP, serta evaluasi menggunakan confusion matrix dengan metrik accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model MLP tanpa ADASYN memperoleh akurasi sebesar 84,65%, sedangkan model MLP dengan ADASYN memperoleh akurasi sebesar 88,83%. Selain itu, nilai precision meningkat dari 75,90% menjadi 88,58%, recall dari 82,89% menjadi 89,51%, dan F1-score dari 79,24% menjadi 89,04%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa penerapan ADASYN efektif dalam mengatasi ketidakseimbangan data dan meningkatkan kinerja algoritma MLP untuk klasifikasi diagnosis penyakit Alzheimer.

Abstract: Alzheimer's disease is a neurodegenerative disorder that is one of the leading causes of dementia and is characterized by progressive cognitive decline. As the number of Alzheimer's patients continues to increase, effective classification methods are needed to support early and accurate diagnosis. One of the major challenges in medical datasets is class imbalance, which can reduce the performance of classification models. This study aims to improve the performance of Alzheimer's disease diagnosis classification by applying the Adaptive Synthetic Sampling (ADASYN) technique and the Multilayer Perceptron (MLP) algorithm. The dataset used in this study was obtained from Kaggle and includes demographic information, health conditions, daily activities, and clinical indicators for patients. The research process includes data preprocessing, feature selection using Information Gain, data balancing using ADASYN, model development using MLP, and evaluation through a confusion matrix with accuracy, precision, recall, and F1-score metrics. The results show that the MLP model without ADASYN achieved 84.65% accuracy, while the MLP

model with ADASYN achieved 88.83% accuracy. Furthermore, precision increased from 75.90% to 88.58%, recall from 82.89% to 89.51%, and F1-score from 79.24% to 89.04%. These findings indicate that ADASYN is effective in addressing class imbalance and significantly improves the performance of the MLP algorithm for Alzheimer's disease diagnosis classification.

Corresponding Author:

Fauzan Azim

Email: fauzanazim17@umri.ac.id

PENDAHULUAN

Penyakit Alzheimer merupakan salah satu gangguan neurodegeneratif yang menjadi penyebab utama demensia pada populasi lanjut usia. Penyakit ini ditandai dengan penurunan fungsi kognitif secara progresif yang memengaruhi kemampuan berpikir, mengingat, berkomunikasi, dan menjalankan aktivitas sehari-hari. Menurut World Health Organization (2023), jumlah penderita demensia di dunia terus meningkat dan diperkirakan akan mencapai 139 juta orang pada tahun 2050. Kondisi tersebut menjadikan Alzheimer sebagai salah satu tantangan utama dalam bidang kesehatan global karena berdampak pada kualitas hidup pasien, keluarga, serta sistem pelayanan kesehatan. Oleh karena itu, diperlukan metode diagnosis yang mampu membantu proses identifikasi penyakit Alzheimer secara lebih cepat dan akurat sehingga penanganan dapat dilakukan sejak dini (Alzheimers, 2024).

Perkembangan teknologi informasi dan kecerdasan buatan telah memberikan kontribusi signifikan dalam bidang kesehatan, khususnya dalam pengembangan sistem pendukung keputusan untuk diagnosis penyakit (Armin et al., 2026; Janna & Lizar, 2026). Salah satu pendekatan yang banyak digunakan adalah machine learning karena mampu mempelajari pola dari data dan menghasilkan prediksi secara otomatis (Muis et al., 2025). Berbagai penelitian menunjukkan bahwa algoritma machine learning dapat digunakan untuk mendeteksi penyakit Alzheimer dengan tingkat akurasi yang cukup tinggi (Hendrastuty et al., 2026; Rani et al., 2024). Salah satu algoritma yang sering digunakan dalam proses klasifikasi adalah Multilayer Perceptron (MLP), yaitu jaringan saraf tiruan yang mampu mempelajari hubungan nonlinier yang kompleks pada data medis (Bhatt et al., 2021). Pemanfaatan teknologi informasi juga telah diterapkan di berbagai bidang untuk meningkatkan efektivitas pengolahan data dan pengambilan keputusan. Maruf et al. (2023) berhasil mengembangkan sistem manajemen bandwidth jaringan berbasis teknologi informasi, sedangkan Sahlan et al., (2026), mengembangkan media pembelajaran digital yang mampu meningkatkan efektivitas proses pembelajaran. Hasil penelitian tersebut menunjukkan bahwa penerapan teknologi informasi dan kecerdasan buatan memiliki potensi besar untuk menyelesaikan berbagai permasalahan, termasuk di bidang kesehatan.

Permasalahan yang umum ditemukan pada dataset medis adalah ketidakseimbangan distribusi kelas (*class imbalance*), yaitu kondisi ketika jumlah data pada suatu kelas jauh lebih banyak dibandingkan kelas lainnya. Pada kasus diagnosis Alzheimer, jumlah data pasien yang teridentifikasi positif sering kali lebih sedikit dibandingkan dengan data pasien normal, sehingga model cenderung menghasilkan prediksi yang bias pada kelas mayoritas. Kondisi tersebut dapat menyebabkan menurunnya kemampuan model dalam mengenali pasien yang benar-benar mengalami Alzheimer. Beberapa penelitian telah mencoba mengatasi permasalahan tersebut dengan menggunakan berbagai teknik *oversampling*, seperti Synthetic Minority Over-sampling Technique (SMOTE) dan Adaptive Synthetic Sampling (ADASYN). Penelitian Yasirah & Wirayuda (2025) menunjukkan bahwa penggunaan teknik *oversampling* dapat meningkatkan performa klasifikasi Alzheimer pada data MRI. Selain itu, penelitian Bagir et al., (2025) dan Sirait et al. (2024) juga membuktikan bahwa ADASYN efektif dalam meningkatkan representasi kelas minoritas sehingga menghasilkan model klasifikasi yang lebih seimbang.

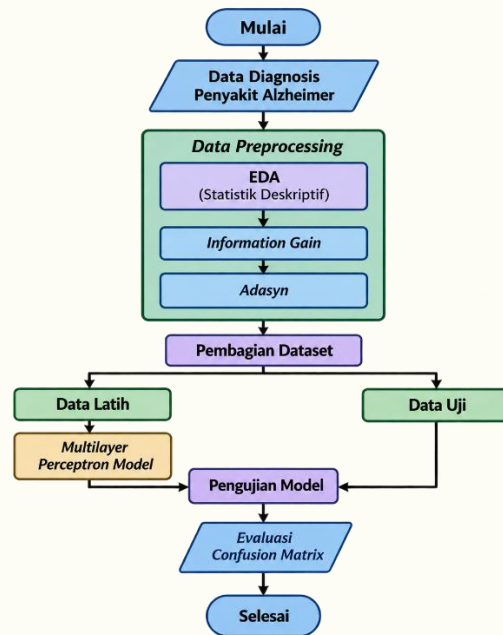
Berdasarkan permasalahan tersebut, penelitian ini mengusulkan penerapan Adaptive Synthetic Sampling (ADASYN) yang dikombinasikan dengan algoritma Multilayer Perceptron untuk klasifikasi diagnosis penyakit Alzheimer. ADASYN digunakan untuk menghasilkan data sintesis pada kelas minoritas berdasarkan tingkat kesulitan setiap sampel sehingga distribusi data menjadi lebih seimbang (Sirait et al., 2024). Selanjutnya, algoritma MLP digunakan sebagai model klasifikasi karena memiliki kemampuan yang baik dalam mempelajari pola kompleks pada data kesehatan (Pratiwi et al., 2024). Proses penelitian meliputi tahap *preprocessing* data, seleksi fitur menggunakan Information Gain, penyeimbangan data menggunakan ADASYN, pembangunan model MLP, serta evaluasi menggunakan *confusion matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

Nilai kebaruan (*novelty*) dari penelitian ini terletak pada kombinasi metode Information Gain, ADASYN, dan Multilayer Perceptron untuk klasifikasi diagnosis Alzheimer menggunakan dataset klinis yang diperoleh dari Kaggle. Berbeda dengan beberapa penelitian sebelumnya yang lebih banyak memanfaatkan data citra MRI atau menggunakan algoritma klasifikasi lain, penelitian ini berfokus pada pemanfaatan data demografi, kondisi kesehatan, aktivitas harian, dan indikator klinis pasien untuk membangun model klasifikasi yang lebih efisien. Selain itu, penelitian ini juga mengevaluasi secara komprehensif pengaruh penerapan ADASYN terhadap peningkatan performa algoritma MLP dalam menangani ketidakseimbangan data. Kombinasi metode yang digunakan diharapkan mampu menghasilkan klasifikasi yang lebih baik dibandingkan dengan penelitian sebelumnya yang hanya menggunakan algoritma klasifikasi tanpa teknik penyeimbangan data.

Penelitian ini bertujuan untuk meningkatkan performa klasifikasi diagnosis penyakit Alzheimer melalui penerapan metode ADASYN dan algoritma Multilayer Perceptron. Hasil penelitian diharapkan dapat memberikan kontribusi dalam pengembangan sistem pendukung keputusan berbasis machine learning di bidang kesehatan, khususnya dalam membantu proses diagnosis dini penyakit Alzheimer. Selain itu, penelitian ini diharapkan dapat menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan klasifikasi data medis menggunakan teknik penyeimbangan data dan jaringan saraf tiruan, serta memperkaya pemanfaatan teknologi informasi dalam penyelesaian permasalahan kesehatan berbasis data.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk mengklasifikasikan diagnosis penyakit Alzheimer menggunakan teknik Adaptive Synthetic Sampling (ADASYN) dan algoritma Multilayer Perceptron (MLP). Metode eksperimen dipilih karena memungkinkan pengukuran secara langsung terhadap pengaruh penerapan ADASYN dalam mengatasi ketidakseimbangan data terhadap kinerja model klasifikasi (Susilo et al., 2025). Tahapan penelitian dilakukan secara sistematis, mulai dari pengumpulan data, prapemrosesan data (Preprocessing), pembangunan model klasifikasi, hingga evaluasi kinerja model menggunakan Confusion Matrix. Alur penelitian terdiri atas pengumpulan dataset, seleksi fitur menggunakan Information Gain, pembagian dataset menjadi data latih dan data uji, seleksi fitur menggunakan Information Gain pada data latih, penyeimbangan data menggunakan ADASYN hanya pada data latih, pembangunan model MLP, serta pengukuran performa model pada data uji.



Gambar 1 Flowchart Metode Penelitian

Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah Alzheimer's Disease Dataset yang diperoleh dari platform Kaggle di alamat <https://www.kaggle.com/datasets/rabieelkharoua/alzheimers-disease-dataset>. Dataset tersebut merupakan hasil kompilasi dari beberapa sumber data publik, termasuk Open Access Series of Imaging Studies (OASIS), yang memuat data demografi, kondisi kesehatan, aktivitas harian, serta indikator klinis pasien yang berkaitan dengan diagnosis penyakit Alzheimer. Dataset digunakan sebagai sumber data utama untuk membangun model klasifikasi diagnosis Alzheimer.

Processing Data

Tahap processing dilakukan untuk meningkatkan kualitas data sebelum digunakan dalam proses klasifikasi. Pada tahap ini dilakukan pemeriksaan data, penanganan missing value, transformasi atribut kategorikal menjadi numerik, serta normalisasi data. Selanjutnya dilakukan seleksi fitur menggunakan metode Information Gain untuk menentukan atribut yang memiliki kontribusi terbesar terhadap diagnosis Alzheimer. Seleksi fitur dilakukan dengan memilih atribut yang memiliki nilai Information Gain tertinggi, sehingga hanya fitur yang relevan digunakan dalam proses klasifikasi. Setelah proses seleksi fitur selesai, dilakukan penyeimbangan distribusi kelas menggunakan metode Adaptive Synthetic Sampling (ADASYN). Metode ini menghasilkan data sintesis pada kelas minoritas berdasarkan tingkat kesulitan setiap sampel sehingga distribusi data menjadi lebih seimbang dan mampu mengurangi bias model terhadap kelas mayoritas.

Pembagian Dataset

Dataset yang telah melalui proses seleksi fitur dan penyeimbangan data selanjutnya dibagi menjadi data latih (training data) dan data uji (testing data). Pembagian data dilakukan menggunakan metode train-test split dengan proporsi 80:20. Sebanyak 1.719 data (80%) digunakan sebagai data latih untuk membangun model klasifikasi, sedangkan 430 data (20%) digunakan sebagai data uji untuk mengevaluasi performa model pada data yang belum pernah dipelajari sebelumnya. Pembagian data dilakukan secara acak dengan mempertahankan proporsi kelas pada masing-masing kelompok data.

Pemodelan Klasifikasi

Tahap pemodelan dilakukan menggunakan algoritma Multilayer Perceptron (MLP) yang merupakan salah satu metode jaringan saraf tiruan (Artificial Neural Network). Implementasi model dilakukan menggunakan pustaka Scikit-Learn pada bahasa pemrograman Python melalui kelas

MLPClassifier. Parameter utama yang digunakan meliputi jumlah neuron pada lapisan tersembunyi (hidden layer), fungsi aktivasi, dan learning rate. Model dilatih menggunakan data latih yang telah melalui proses seleksi fitur dan penyeimbangan data menggunakan ADASYN. Setelah proses pelatihan selesai, model digunakan untuk melakukan prediksi terhadap data uji guna memperoleh hasil klasifikasi diagnosis penyakit Alzheimer.

Evaluasi Model

Kinerja model dievaluasi menggunakan Confusion Matrix dengan membandingkan hasil prediksi model dengan label sebenarnya pada data uji. Evaluasi dilakukan menggunakan empat metrik utama, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Metrik tersebut digunakan untuk mengukur tingkat ketepatan model secara keseluruhan, kemampuan model dalam mendeteksi kasus positif Alzheimer, serta keseimbangan antara nilai *precision* dan *recall*. Penggunaan beberapa metrik evaluasi diperlukan karena dataset memiliki karakteristik ketidakseimbangan kelas, sehingga pengukuran performa tidak cukup hanya menggunakan nilai akurasi.

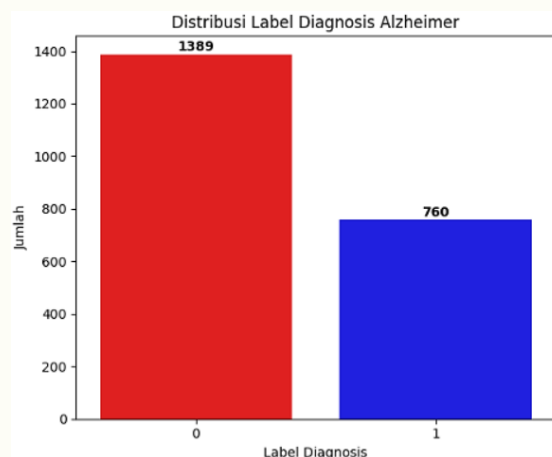
Teknik Pengumpulan Data

Analisis data dilakukan dengan membandingkan performa model Multilayer Perceptron sebelum dan sesudah penerapan ADASYN. Nilai *accuracy*, *precision*, *recall*, dan *F1-score* digunakan sebagai dasar untuk menentukan efektivitas metode yang diusulkan. Model dengan nilai evaluasi yang lebih tinggi dianggap memiliki kemampuan klasifikasi yang lebih baik dalam mendeteksi diagnosis penyakit Alzheimer. Kesimpulan penelitian ditarik berdasarkan hasil perbandingan kedua model untuk mengetahui kontribusi ADASYN dalam meningkatkan kinerja algoritma Multilayer Perceptron pada klasifikasi diagnosis penyakit Alzheimer.

HASIL DAN PEMBAHASAN

Analisis dan Visualisasi Dataset Diagnosis Penyakit Alzheimer

Dataset yang digunakan dalam penelitian ini adalah Alzheimer's Disease Dataset yang diperoleh dari platform Kaggle. Dataset terdiri atas 2.149 data pasien dengan 35 atribut yang mencakup informasi demografi, kondisi kesehatan, aktivitas harian, serta indikator klinis yang berkaitan dengan diagnosis penyakit Alzheimer. Variabel target yang digunakan adalah atribut Diagnosis yang terdiri atas dua kelas, yaitu positif Alzheimer dan negatif Alzheimer. Dataset ini dipilih karena memiliki atribut yang cukup lengkap untuk menggambarkan kondisi pasien secara komprehensif dalam proses klasifikasi diagnosis penyakit Alzheimer.



Gambar 1. Distribusi Jumlah Label Diagnosis Penyakit Alzheimer

Pada Gambar 4.1 di atas, sumbu horizontal (x) menunjukkan dua kategori label diagnosis, yaitu 0 dan 1, yang masing-masing merepresentasikan kondisi pasien. Label 0 digunakan untuk menunjukkan pasien dengan kondisi normal atau tidak menderita Alzheimer, sedangkan label 1

mengindikasikan bahwa pasien telah terdiagnosis Alzheimer. Sumbu vertikal (y) menampilkan jumlah sampel atau kasus untuk setiap label. Grafik batang pada Gambar 4.2 di atas memperlihatkan bahwa jumlah data dengan label 0 sebanyak 1.389 kasus, sedangkan label 1 sebanyak 760 kasus. Hal ini menunjukkan adanya ketidakseimbangan kelas (class imbalance) dalam dataset yang digunakan, karena berdasarkan dataset tersebut, jumlah data pasien yang tidak mengidap Alzheimer jauh lebih banyak dibandingkan dengan pasien yang mengidap Alzheimer. Ketidakseimbangan ini dapat memengaruhi performa model pembelajaran mesin yang akan dilatih.

Selain distribusi kelas, dilakukan visualisasi terhadap atribut usia, jenis kelamin, dan aktivitas harian pasien untuk memahami karakteristik data yang digunakan. Hasil visualisasi menunjukkan bahwa sebagian besar pasien berada pada kelompok usia lanjut yang merupakan kelompok risiko utama penyakit Alzheimer. Temuan ini sesuai dengan penelitian Digambiro et al. (2024) yang menyatakan bahwa peningkatan usia merupakan faktor risiko dominan dalam perkembangan penyakit Alzheimer.

Pra-pemrosesan Dataset Diagnosis Penyakit Alzheimer

Tahap pra-pemrosesan dilakukan untuk meningkatkan kualitas data sebelum digunakan dalam proses klasifikasi. Tahapan yang dilakukan meliputi pemeriksaan missing value, seleksi fitur menggunakan Information Gain, serta penyeimbangan data menggunakan ADASYN. Pemeriksaan missing value dilakukan pada seluruh atribut dalam dataset.

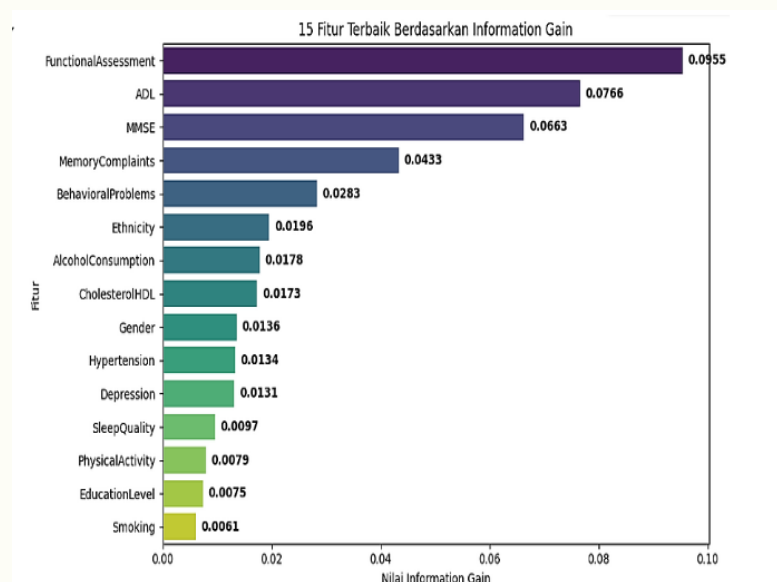
Tabel 1. Statistik Missing Value Dataset Diagnosis Penyakit Alzheimer

Keterangan	Nilai
Jumlah Data	2.149
Jumlah Atribut	35
Missing Value	0

Hasil pemeriksaan menunjukkan bahwa seluruh atribut tidak memiliki data yang hilang, sehingga seluruh data dapat langsung digunakan dalam proses selanjutnya tanpa memerlukan teknik imputasi data. Kondisi ini menunjukkan bahwa kualitas dataset yang digunakan cukup baik untuk proses klasifikasi.

Seleksi Fitur Menggunakan Information Gain

Seleksi fitur dilakukan menggunakan metode Information Gain untuk menentukan atribut yang memiliki kontribusi terbesar terhadap diagnosis penyakit Alzheimer.

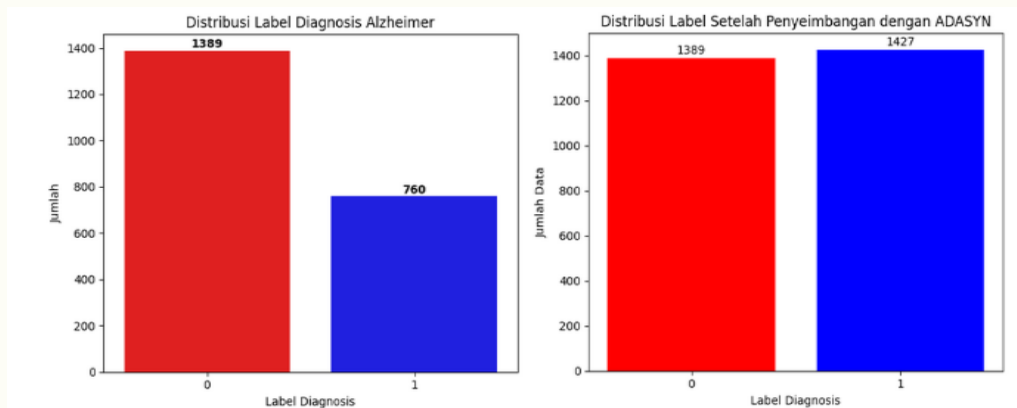


Gambar 2. Hasil Pemilihan 15 Fitur Dengan Information Gain

Gambar di atas menunjukkan hasil seleksi 15 fitur menggunakan metode Information Gain yang akan digunakan untuk membangun model klasifikasi diagnosis penyakit Alzheimer. Pemilihan 15 fitur dilakukan dengan alasan, agar dapat mempertahankan informasi antarfitur pada suatu data, maka setidaknya menggunakan fitur sebanyak kurang dari 50% dari total 35 fitur yang ada pada data diagnosis penyakit Alzheimer. Fitur dengan nilai information gain tertinggi dianggap paling relevan untuk membedakan individu yang positif dan negatif terhadap penyakit tersebut. Atribut dengan nilai information gain tertinggi adalah FunctionalAssessment dengan skor 0.0955. Hasil seleksi fitur menunjukkan bahwa atribut MMSE, Functional Assessment, dan ADL merupakan fitur yang memiliki kontribusi terbesar terhadap diagnosis penyakit Alzheimer. Temuan ini sejalan dengan penelitian Rani et al. (2024) yang menyatakan bahwa skor kognitif dan kemampuan aktivitas harian merupakan indikator penting dalam identifikasi penyakit Alzheimer.

Penyeimbangan Data Menggunakan ADASYN

Setelah proses seleksi fitur, dilakukan penyeimbangan data menggunakan metode Adaptive Synthetic Sampling (ADASYN). Metode ini digunakan untuk mengatasi ketidakseimbangan distribusi kelas dengan membangkitkan data sintetis untuk kelas minoritas.

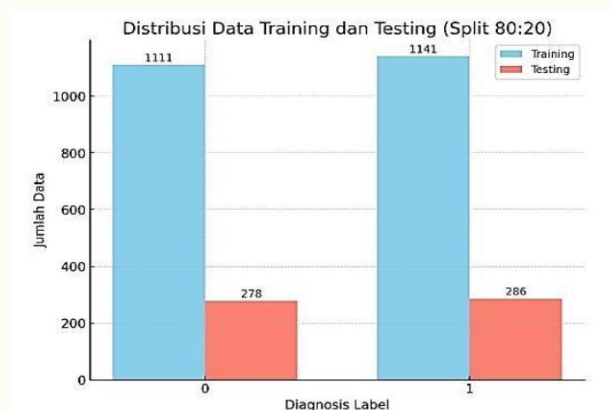


Gambar 3. Distribusi Label Diagnosis Penyakit Alzheimer Setelah ADASYN

Hasil penerapan ADASYN menunjukkan bahwa jumlah data pada kedua kelas menjadi lebih seimbang dibandingkan dengan kondisi awal dataset. Dengan distribusi data yang lebih seimbang, model dapat mempelajari karakteristik setiap kelas secara lebih optimal sehingga mengurangi bias terhadap kelas mayoritas.

Pembagian Dataset

Dataset yang telah melalui proses seleksi fitur dan penyeimbangan data kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20.



Gambar 4. Split Dataset 80:20 Diagnosis Penyakit Alzheimer

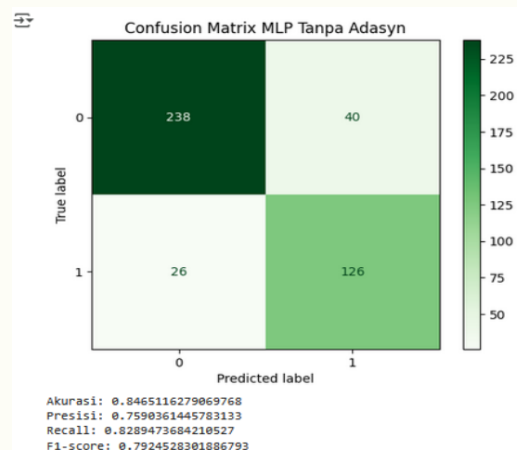
Gambar di atas menampilkan grafik batang yang menunjukkan distribusi label diagnosis penyakit Alzheimer setelah dilakukan proses pembagian data (split data) dengan skenario rasio 80:20, di mana 80% data digunakan sebagai data latih (training) dan 20% sebagai data uji (testing). Label 0 merepresentasikan pasien yang tidak terdiagnosis Alzheimer, sementara label 1 merepresentasikan pasien yang terdiagnosis Alzheimer. Berdasarkan grafik, data latih (ditampilkan dalam batang berwarna biru muda) menunjukkan bahwa terdapat 1111 data berlabel 0 dan 1141 data berlabel 1. Ini mengindikasikan bahwa distribusi label pada data latih cukup seimbang, yang sangat penting untuk memastikan bahwa model pembelajaran mesin dapat mempelajari karakteristik kedua kelas secara proporsional. Sementara itu, data uji (ditampilkan dalam batang berwarna merah muda) terdiri dari 278 data berlabel 0 dan 286 data berlabel 1. Hal ini menunjukkan bahwa pembagian data uji juga dilakukan secara proporsional terhadap distribusi data latih. Dengan distribusi data yang seimbang pada kedua subset data ini, model diharapkan memiliki kinerja yang lebih adil dan tidak bias terhadap salah satu kelas diagnosis. Pembagian dataset dilakukan menggunakan metode train-split untuk memastikan model dapat diuji menggunakan data yang belum pernah dipelajari selama proses pelatihan.

Pemodelan Klasifikasi

Tahap pemodelan dilakukan menggunakan algoritma Multilayer Perceptron Classifier (MLP). Algoritma ini dipilih karena mampu mempelajari hubungan nonlinier antarfitur dan telah banyak digunakan dalam berbagai penelitian klasifikasi data medis. Model dibangun menggunakan data latih yang telah melalui proses seleksi fitur dan penyeimbangan data. Model MLP terdiri atas lapisan input, lapisan tersembunyi (hidden layer), dan lapisan output yang digunakan untuk menghasilkan prediksi diagnosis penyakit Alzheimer.

Pengujian dan Evaluasi Model

Skenario 1: Pengujian pertama dilakukan menggunakan algoritma MLP tanpa penerapan ADASYN.

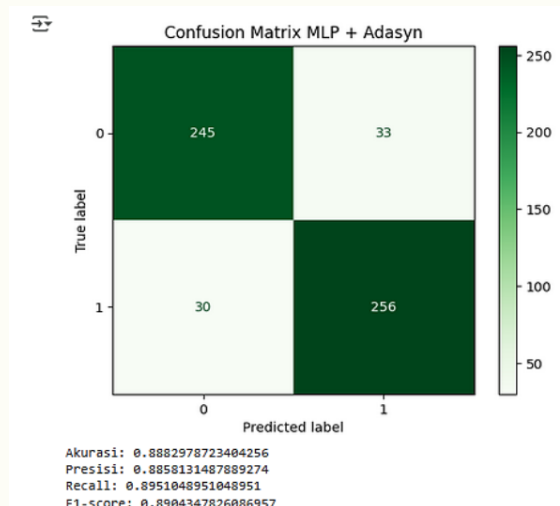


Gambar 6. Confusion Matrix MLP Tanpa ADASYN

Gambaran menyeluruh mengenai performa algoritma Multilayer Perceptron (MLP) dalam melakukan klasifikasi diagnosis penyakit Alzheimer, tanpa melalui proses penyeimbangan data menggunakan teknik Adaptive Synthetic Sampling (ADASYN). Pada gambar tersebut, terdapat dua kelas, yakni label 0 dan label 1, yang masing-masing merepresentasikan kategori diagnosis tertentu (misalnya, 0 = non-Alzheimer, 1 = Alzheimer). Confusion matrix ini menyajikan jumlah klasifikasi yang benar dan salah untuk setiap kelas berdasarkan data uji. Dari hasil tersebut, diperoleh:

1. Sebanyak 238 sampel kelas 0 berhasil diklasifikasikan dengan benar sebagai kelas 0 (True Negative),
2. Sebanyak 126 sampel kelas 1 berhasil diklasifikasikan dengan benar sebagai kelas 1 (True Positive),
3. Terdapat 40 sampel kelas 0 yang salah diklasifikasikan sebagai kelas 1 (False Positive),
4. Sebanyak 26 sampel kelas 1 salah diklasifikasikan sebagai kelas 0 (false negative).

Hasil evaluasi menunjukkan bahwa model memperoleh nilai akurasi sebesar **84,65%**, precision **75,90%**, recall **82,89%**, dan F1-score **79,24%**. Hasil tersebut menunjukkan bahwa model masih mengalami kesulitan dalam mengenali kelas minoritas secara optimal. Skenario 2: Pengujian kedua dilakukan menggunakan kombinasi Information Gain, ADASYN, dan MLP.



Gambar 7. Confusion Matrix MLP Dengan ADASYN

Confusion Matrix di atas menampilkan hasil evaluasi performa model Multilayer Perceptron (MLP) dengan penerapan teknik Adaptive Synthetic Sampling (ADASYN) pada tugas klasifikasi diagnosis penyakit Alzheimer. Pada gambar tersebut, terdapat dua kelas, yakni label 0 dan label 1, yang masing-masing merepresentasikan kategori diagnosis tertentu (misalnya, 0 = non-Alzheimer, 1 = Alzheimer). Nilai-nilai dalam matriks ini menunjukkan hasil klasifikasi terhadap data uji (testing data), dan digunakan untuk menghitung metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Dalam confusion matrix tersebut, dapat diuraikan masing-masing nilai sebagai berikut:

1. 245 sampel kelas 0 berhasil diklasifikasikan dengan benar sebagai kelas 0 (True Negative),
2. 256 sampel kelas 1 berhasil diklasifikasikan dengan benar sebagai kelas 1 (True Positive),
3. 33 sampel kelas 0 diklasifikasikan secara salah sebagai kelas 1 (False Positive),
4. 30 sampel kelas 1 diklasifikasikan secara salah sebagai kelas 0 (False Negative).

Hasil evaluasi menunjukkan peningkatan performa pada seluruh metrik evaluasi. Model memperoleh nilai accuracy sebesar **88,83%**, precision **88,58%**, recall **89,51%**, dan F1-score **89,04%**. Hasil tersebut menunjukkan bahwa ADASYN mampu meningkatkan kemampuan model dalam mengenali kelas positif Alzheimer.

Analisis Performa Model MLP

Perbandingan hasil evaluasi kedua model disajikan pada Tabel 4.

Tabel 2. Perbandingan Evaluasi Model MLP

Metrik	MLP Tanpa ADASYN	MLP + ADASYN
Accuracy	84,65%	88,83%
Precision	75,90%	88,58%
Recall	82,89%	89,51%
F1-Score	79,24%	89,04%

Berdasarkan Tabel 4, penerapan ADASYN meningkatkan akurasi sebesar **4,18%**, precision sebesar **12,68%**, recall sebesar **6,62%**, dan F1-score sebesar **9,80%**. Peningkatan tersebut menunjukkan bahwa ADASYN berhasil mengatasi ketidakseimbangan data sehingga model dapat mengenali kelas minoritas dengan lebih baik.

Hasil penelitian ini sejalan dengan penelitian Yasirah & Wirayuda (2025) yang menunjukkan bahwa teknik oversampling mampu meningkatkan kinerja klasifikasi Alzheimer. Dengan demikian, kombinasi ADASYN dan Multilayer Perceptron terbukti efektif untuk meningkatkan akurasi klasifikasi diagnosis penyakit Alzheimer berbasis data klinis.

KESIMPULAN

Penelitian ini bertujuan untuk meningkatkan performa klasifikasi diagnosis penyakit Alzheimer melalui penerapan metode Adaptive Synthetic Sampling (ADASYN) dan algoritma Multilayer Perceptron (MLP) pada dataset yang memiliki ketidakseimbangan distribusi kelas. Berdasarkan hasil penelitian, kombinasi kedua metode tersebut terbukti mampu menghasilkan model klasifikasi yang lebih efektif dalam mengenali pola data pasien Alzheimer dibandingkan dengan penggunaan algoritma MLP tanpa proses penyeimbangan data. Penerapan ADASYN memberikan kontribusi dalam meningkatkan representasi kelas minoritas sehingga proses pembelajaran model menjadi lebih seimbang dan mampu menghasilkan prediksi yang lebih akurat.

Hasil penelitian menunjukkan bahwa penanganan ketidakseimbangan data merupakan faktor penting dalam pengembangan model klasifikasi di bidang kesehatan, khususnya untuk diagnosis penyakit Alzheimer. Penggunaan seleksi fitur Information Gain juga membantu model memfokuskan proses pembelajaran pada atribut yang paling relevan, sehingga meningkatkan efisiensi dan kualitas klasifikasi. Dengan demikian, pendekatan yang diusulkan dapat menjadi alternatif potensial untuk mendukung sistem diagnosis dini Alzheimer berbasis machine learning.

Penelitian ini memiliki prospek untuk dikembangkan lebih lanjut melalui penggunaan dataset yang lebih besar dan beragam, sehingga model yang dihasilkan memiliki kemampuan generalisasi yang lebih baik. Selain itu, penelitian selanjutnya dapat membandingkan performa ADASYN dengan metode penyeimbangan data lainnya atau mengombinasikannya dengan algoritma klasifikasi berbasis *deep learning* untuk memperoleh hasil yang lebih optimal. Dari sisi implementasi, model yang dikembangkan berpotensi diterapkan sebagai bagian dari sistem pendukung keputusan pada layanan kesehatan untuk membantu tenaga medis dalam melakukan identifikasi dini penyakit Alzheimer secara lebih cepat, objektif, dan akurat.

DAFTAR PUSTAKA

- Alzheimers, A. (2024). 2024 Alzheimers Disease Facts And Figures. *Alzheimers Dementia*, 20(5), 3708–3821. <https://doi.org/10.1002/alz.13809>
- Armin, M., Arifin, N., & Rachmini, S. A. (2026). Penerapan Algoritma Random Forest dengan Seleksi Fitur Recursive Feature Elimination (RFE) untuk Klasifikasi Gagal Jantung. *Jurnal Sistem Komputer dan Kecerdasan Buatan*, 9(3), 303-311. <https://doi.org/10.47970/siskom-kb.v9i3.975>
- Bagir, M., Mayatopani, H., Riyanto, U., & Alamsyah, D. (2025). Klasifikasi Risiko Diabetes Mellitus Menggunakan K-Nearest Neighbors dengan Peningkatan Performa Melalui Teknik Oversampling ADASYN. *Journal of Information System Research (JOSH)*, 6(4), 2238-2247. <https://doi.org/10.47065/josh.v6i4.7237>
- Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K., & Ghayvat, H. (2021). CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope. *Electronic*, 10(20), 2470. <https://doi.org/10.3390/electronics10202470>
- Hendrastuty, N., Wijaya, M. S., & Alita D. (2026) Deteksi Alzheimer Berbasis Random Forest dengan Feature Selection dan Analisis SHAP pada Data Klinis. *Journal of Information Technology, Software Engineering and Computer Science (ITSECS)*, 4(4), 153-171. <https://doi.org/10.58602/itsecs.v4i4.405>

- Janna, N., & Lizar, Y. (2026). Implementasi Metode Forward Chaining pada Sistem Pakar untuk Diagnosa Penyakit sebagai Pendukung Keputusan. *Indonesian Journal of Multidisciplinary on Social and Technology*, 4(2), 1296-1302. <https://doi.org/10.69693/ijmst.v4i2.9424>
- Maruf, M. A. A., Darman, D., & Razilu, Z. (2023). Rancang Bangun Manajemen Bandwidth Jaringan Pada Laboratorium Teknik Komputer Dan Jaringan. *Decode :Jurnal Pendidikan Teknologi Informasi*, 3(2), 246–256. <https://doi.org/10.51454/decode.v3i2.177>
- Muis, A., Rustiawan, A., Oyeyemi, B. B., Syukur, A., & Furizal. (2025). Early Detection of Brain Tumors: Performance Evaluation of AlexNet and GoogleNet on Different Medical Image Resolutions. *Scientific Journal of Engineering Research*, 1(3), 120–130. <https://doi.org/10.64539/sjer.v1i3.2025.10>
- Pratiwi, N. K. C., Ibrahim, N., & Saidah, S. (2024). Prediksi Kanker Paru menggunakan Grid search untuk Optimasi Hyperparameter pada Algoritma MLP dan Logistic Regression. *ELKOMIK*, 12(3), 556-568. <http://dx.doi.org/10.26760/elkomika.v12i3.556>
- Rani, P., Lamba, R., Sachdeva, R. K., Kumar, K., & Iwendi, C. (2024). A Machine Learning Model For Alzheimer's Disease Prediction. *IET Cyber-Physical Systems: Theory & Applications*, 9(2). <https://doi.org/10.1049/cps2.12090>
- Sahlan, M., Fitrianiingsih, N., Fitriati, I., Diana, N., & Hardiansyah, H. (2026). Media Gamifikasi WordWall dan QuizWhizzer untuk Meningkatkan Kolaborasi Siswa. *Decode: Jurnal Pendidikan Teknologi Informasi*, 6(2), 557–569. <https://doi.org/10.51454/decode.v6i2.1742>
- Sirait, M. T. T. B., Fathonah, N. S. & Fauzan, M. N. (2024). Pemanfaatan Algoritma Adasyn Dan Support Vector Machine Dalam Meningkatkan Akurasi Prediksi Kanker Paru-Paru. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(5), 8773-8778. <https://doi.org/10.36040/jati.v8i5.10752>
- Susilo, L. B. E., Siswa, T. A. Y., & Verdikha, N. A. (2025). Penerapan Algoritma Decision Tree, Recursive Elimination Dan ADASYN Pada Data Klasifikasi Penyakit Dermatitis. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(5), 8792-8798. <https://doi.org/10.36040/jati.v9i5.15212>
- World Health Organization. (2023). *Dementia*. <https://www.who.int/news-room/fact-sheets/detail/dementia>
- Yasirah, F., & Wirayuda, T. A. B. (2025). Classification of Alzheimer's Disease based on MRI Image using Convolutional Neural Network. 2025 *International Conference on Data Science and Its Applications (ICoDSA)*, Jakarta, Indonesia, 468-473. <https://doi.org/10.1109/ICoDSA67155.2025.11157551>.