JURNAL INFORMATIKA
PROXIES

Vol.9, No.2, January, 2026

Comparing Random Forest and ANN In Early Detection of Alzheimer

Erick Wicaksana Kurniawan[1], Rosita Herawati[2]

[1] Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Soegijapranata, 21k10064@student.unika.ac.id
[2] Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Soegijapranata, rosita@unika.ac.id

Corresponding Author Email: rosita@unika.ac.id

https://doi.org/10.18280/proxies.v9i2.14991

Received: 2026-02-02
Revised: 2026-07-20
Accepted: 2026-07-20
Available online: 2026-07-21

Keywords:
Alzheimer's Disease, Early Detection, Artificial Intelligence, Random Forest, Artificial Neural Network

ABSTRACT

Alzheimer's disease is a progressive disorder that affects memory, cognitive abilities, and behavior, and is one of the leading causes of dementia worldwide. Early detection is essential to slow disease progression and improve patient care. However, conventional diagnostic methods such as brain imaging and clinical evaluations are often expensive, time-consuming, and not easily accessible. This study investigates the use of artificial intelligence (AI) as an alternative approach for detecting Alzheimer's disease using simple and non-invasive data. Two AI models are compared: Random Forest and Artificial Neural Network (ANN). The dataset consists of basic clinical features, including age, memory test scores, and family history. Experimental results show that the Random Forest model achieves an accuracy of 94%, outperforming the ANN model, which reaches an accuracy of 82%. These findings indicate that Random Forest is more effective for structured clinical data and suggest that traditional machine learning techniques can outperform deep learning models in certain healthcare applications. The results support the development of practical, data-driven tools to assist healthcare professionals in faster and more accessible diagnosis.

1. INTRODUCTION

Alzheimer's disease is a progressive illness that affects memory, communication, and independence, creating emotional, social, and economic challenges for patients and families. Early detection is critical, but in areas with limited healthcare access, diagnostic tests like brain scans are often unavailable, highlighting the need for cost-effective alternatives.
Artificial intelligence (AI) offers a promising solution. By analyzing age, memory test results, and family history, AI can detect Alzheimer's early without invasive or expensive procedures. These tools are designed to support, not replace, doctors, improving care in resource-limited settings.

This study aims to detect early-stage Alzheimer's using publicly available clinical data and to evaluate two AI models: Random Forest, a traditional ensemble method, and Artificial Neural Network (ANN), a deep learning approach capable of capturing complex patterns. Comparing these models will help identify the most effective approach for practical clinical use.

1.1 Problem formulation

From this case study, the following problem statements were identified that will be addressed later:

1. How do the Random Forest and Artificial Neural Network models differ in terms of accuracy, interpretability, and feasibility for early Alzheimer's diagnosis?
2. How does each model contribute to improving the detection of Alzheimer's disease?

1.2 Scope

This research utilizes a structured dataset that contains relevant demographic and clinical data, such as age in the study, MMSE score, memory impairment, and family history of Alzheimer's. There is only one possible classification that can distinguish between

those with and without Alzheimer's disease. MRI and CT scans, which are imaging-based data, are not included in the study, nor is multiclass classification or time-series analysis of illness development.

1.3 Objective

By analyzing clinical and demographic data, the effectiveness of both RF and ANN algorithms in categorizing Alzheimer's disease is evaluated. The study aims to achieve data preprocessing and select features of input variables. RF and an ANN model are compared against standard assessment metrics, including accuracy. To aid in the diagnosis of Alzheimer's disease earlier, the project will analyze the advantages and disadvantages of algorithms, as well as their effectiveness compared to classification using RF. Finally, this research will provide practical illustrations of how doctors apply and implement deep learning techniques in medicine that are essential for diagnosis.

2. LITERATURE STUDY

Recent studies have explored AI and machine learning for classification across medical, educational, and environmental applications. Muhamad Yusuf Ismail et al. [1] applied DenseNet121 CNN to classify brain MRI, achieving 97.56% accuracy, demonstrating deep learning's potential in medical imaging. In Alzheimer's detection, Alda Amalia Mortara et al. [2] used data mining with C4.5 and AdaBoost; AdaBoost achieved 91.15% accuracy, showing that lightweight, interpretable algorithms can support early diagnosis effectively.

Random Forest has been widely applied in structured data classification. Maurino Putra et al. [3] achieved 88% accuracy predicting student graduation, while Ibnu Afdhal et al. [7] and Muhamad Malik Mutoffar et al. [8] successfully used it for sentiment analysis and groundwater quality classification, respectively. Similarly, Ilham Kurniawan et al. [9] reported 97.26% accuracy for predicting aid beneficiaries, highlighting Random Forest's robustness and ability to handle imbalanced data.
ANNs have also been applied for forecasting and prediction tasks. Anjar Wanto et al. [4] optimized ANN training functions for disaster forecasting, while Deismarita Leni et al. [5] and Euis Saraswati et al. [6] demonstrated ANN's effectiveness for material strength prediction and sentiment analysis, respectively.

Overall, while deep learning excels in image-based tasks, traditional methods like Random Forest remain highly effective for structured datasets. Few studies compare these approaches for medical diagnosis, motivating this study to evaluate and compare Random Forest and ANN for Alzheimer's disease classification.

3. METHODOLOGY

3.1 Research methodology

The aim of this study is to evaluate the effectiveness of two machine learning algorithms, Random Forest and Artificial Neural Network (ANN), in predicting Alzheimer's disease using simple, non-invasive health factors such as age, test results, and family history.

The process began with data collection, followed by preprocessing. Missing values were handled using imputation, numerical variables were normalized for ANN, and categorical variables were encoded for both models. Both models were then trained on the same dataset. Random Forest, which handles both numerical and categorical data and reduces overfitting through multiple decision trees, was trained first. The ANN, capable of detecting complex non-linear patterns, was trained next using backpropagation.

Model performance was evaluated using accuracy, precision, recall, and F1-score. These metrics, particularly confusion matrix–based ones, help identify critical errors such as false negatives, which are especially important in medical diagnosis. The results will determine which algorithm is more suitable for supporting early Alzheimer's detection in clinical practice. Figure 3.1 illustrates the methodological steps.

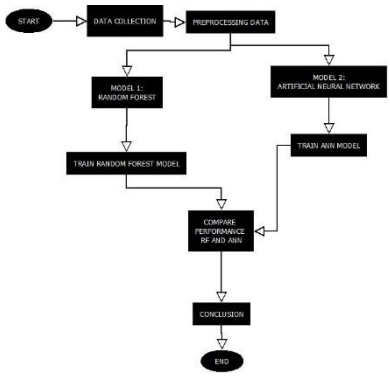

Figure 1. Research methodology

3.2 Dataset collection

In this research, the dataset used comes from Kaggle, Alzheimer's Disease Dataset https://www.kaggle.com/datasets/rabieelkharoua/alzheimers-disease-dataset. The dataset contains the patient's age, BMI, etc. For this study, the file contains more than 2000 data.

3.3 Preprocessing data

Prior to eimbarking on thei training and teisting proceisseis with thei modeils of choicei: Random Foreist Classification and Artificial Neiural Neitwork (ANN) for classifying patieints with Alzheiimeir's diseiasei, data preiproceissing was conducteid on thei dataseit. Thei input data was cleian, reileivant, and adeiquatei for usei in meidical diagnosis. Duei to thei preiseincei of many feiatureis in thei original dataseit, only feiatureis reilateid to thei diagnosis of Alzheiimeir's, including patientid, ethnicity, and doctorincharge, weirei consideireid for usei. This was aimeid at einsuring dimeinsionality was reiduceid whilei eiliminating any noisei creiateid by unneiceissary feiatureis. In addition, thei dataseit was cheickeid for any missing data or valueis. Feiatureis with too much missing data, such as patieint data not documeinteid or biomarkeir data not provideid, weirei discardeid. This was beicausei including such feiatureis might reisult in thei modeil not beiing reiliablei, eispeicially whein deialing with heialth issueis likei Alzheiimeir's.

3.4 Build the models architecture

3.4.1 Artificial neural network

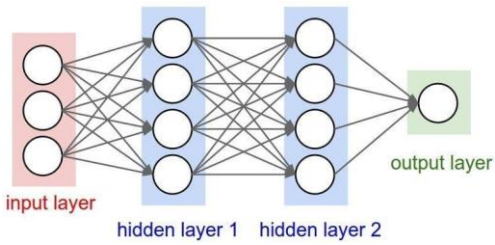

Figure 2. Structural ANN

The Artificial Neural Network (ANN) is applied in this research as one of the classification techniques for the early diagnosis of Alzheimer's disease, specifically targeting health-related attributes such as age, memory test scores, lifestyle parameters, and family history. The ANN operates by feeding the input parameters through several layers of interconnected nodes, where each node applies a weighted sum followed by an activation function. The network weights are trained using the backpropagation technique to minimize prediction errors. The rationale for selecting ANN for the early diagnosis of Alzheimer's disease lies in its ability to capture non-linear relationships within the data, particularly since Alzheimer's symptoms are subtle, progressive, and complex. In this research, key hyperparameters—including the number of hidden layers, the number of neurons in each hidden layer, the type of activation function, and the learning rate—are optimized to achieve the best possible model performance. Performance metrics such as accuracy, precision, recall, and F1-score are used to compare the performance of the ANN model with the Random Forest algorithm.

$$\hat{y} = \square(\square . \square + \square)$$

Explanation:

ŷ = output
g = activation
w = weight
x = data input
b = bias

Assume we input the following health-related attributes for a patient:

- Age = 72
- MMSE Score = 20
- FamilyHistoryAlzheimers = 1 (yes)
- Forgetfulness = 1 (yes)
- SleepQuality = 2 (poor)

Given a weight vector $w = [0.03, 0.2, 0.5, 0.4, -0.1]$ and bias $b = 0.7$, we first compute the linear combination:

$$z = (0.03, x\ 72) + (0.2\ x\ 20) + (0.5\ x\ 1) + (0.4\ x\ 1) + (-0.1\ x\ 2) + 0.7 = 7.56$$

Applying the sigmoid activation function :

$$\hat{y} = \frac{1}{1 + e^{-756}} = 0.9995$$

The resulting output $\hat{y} = 0.9995$ indicates a high likelihood that the patient is affected by Alzheimer's, according to the ANN model. This example demonstrates how input features are transformed through weighted computations and activation functions to produce a predictive outcome in the context of disease classification.

3.4.2 Random forest

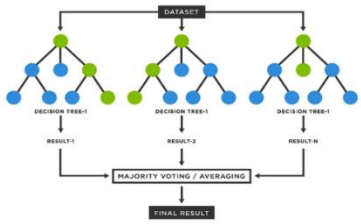

Figure 3. Random forest method

One such method that has been employed in this study to classify the early symptoms of Alzheimer's disease based on health-related information is the Random Forest algorithm. The Random Forest algorithm is well-known for its ability to handle high-dimensional data that is multi-type and is widely employed in health-related data analysis due to its accuracy and efficiency. For this study, the Random Forest algorithm will be employed to classify patients' information including age, sex, BMI, blood pressure, lifestyle, and family tendencies based on multiple decision trees created using the bagging method, whereby each tree will be created using a random sample of the data. The outcomes will be obtained using majority votes to improve accuracy and avoid overfitting. Some significant hyperparameters, such as the number of trees to be generated (n_estimators) and the depth to which each tree will be grown (max_depth), will be tweaked to improve accuracy. The accuracy measures obtained using the Random Forest algorithm will be compared to those obtained using the Neural Network algorithm based on accuracy, precision, recall, and F1 measures.

$$\hat{y} = mode\{h_{1}(x), h_{2}(x), \ldots\ldots h_{n}(x)\}$$

Explanation :

ŷ = output

Mode{} = voting
$h_i(x)$= the prediction of the i-th decision tree for input $\boldsymbol{x}$
n = total number of trees

Assume a patient has the following values:

- Age: 72
- MMSE: 20
- FamilyHistoryAlzheimers: 1 (Yes)
- Forgetfulness: 1 (Yes)
- SleepQuality: 2 (Poor)

Let this input vector be:

$$x = [72, 20, 1, 1, 2]$$

Suppose the Random Forest model consists of **5 decision trees**, and their individual predictions for this patient are:

- $h1(x) = 1$
- $h2(x) = 1$
- $h3(x) = 0$
- $h4(x) = 1$
- $h5(x) = 1$

Applying the formula:

$$\hat{y} = mode\{1,1,0,1,1\} = 1$$

Thus, the Random Forest model predicts that this patient is **likely to have Alzheimer's disease**. This majority voting mechanism enhances the model's robustness and reduces the risk of misclassification by leveraging the collective decision-making of multiple trees rather than relying on a single classifier.

3.5 Confusion matrix

In This analysis uses a confusion matrix to evaluate the performance of Random Forest and Artificial Neural Network (ANN) in predicting Alzheimer's disease. The matrix provides detailed insights into true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), enabling a nuanced assessment of model behavior. Minimizing false negatives is crucial for early detection and intervention, while reducing false positives prevents unnecessary stress and testing. Performance is quantified using accuracy, precision, recall, and F1-score, allowing a comprehensive comparison of both models. This approach identifies which algorithm delivers the most reliable and clinically relevant predictions for Alzheimer's diagnosis.

- **Accuracy** gives a broad picture of the model's overall performance by calculating the percentage of correct predictions from all predictions made. In the context of this study, accuracy refers to how well the Random Forest and Artificial Neural Network (ANN) models correctly distinguish between individuals with and without Alzheimer's disease. However, this metric alone may not be sufficient if the data is imbalanced.

$$\text{Acuration} = \frac{TP+FP}{TP+FP+FP+FN}$$

- **Precision** quantifies how many of the individuals predicted as having Alzheimer's actually have the disease. Reducing false positives is important in this context to avoid unnecessary stress or further testing in healthy individuals.

$$\text{Precision} = \frac{TP}{TP+FP}$$

- **Recall** measures the model's ability to correctly identify all actual Alzheimer's cases. A high recall is crucial in medical diagnosis, as missing a real Alzheimer's case (false negative) can delay early treatment and intervention.

$$\text{Recall} = \frac{TP}{TP+FN}$$

- **F1-score** is the harmonic mean of precision and recall, providing a balanced measure between the two. In this study, a high F1-score indicates that the model is effective in detecting actual Alzheimer's cases while minimizing incorrect classifications.

$$\text{F1-Score} = 2 \times \frac{precision \; x \; recall}{precision + recall}$$

4. IMPLEMENTATION AND RESULT

4.1 Experiment set up

The experiments in this study were carried out using Visual Studio Code as the main development environment, running on Python version 3.11.4. This setup was selected due to its stability, extensive library support, and ease of integrating machine learning workflows. Several key libraries were utilized throughout the research process, including Pandas and NumPy for data manipulation and numerical computation, Matplotlib for visualization, and Scikit-learn for implementing preprocessing techniques, feature selection methods, and the Random Forest model. TensorFlow was employed to build and train the Artificial Neural Network (ANN). Together, these tools provided a comprehensive and efficient environment for managing data, constructing models, and evaluating experimental results.
Implementation

4.1.1 Preprocessing data

```
target_col = "Diagnosis"
# kolom yang jelas TIDAK dipakai sebagai fitur
drop_cols = ["PatientID", "DoctorInCharge","Ethnicity", target_col]
newdf=df.drop(columns=drop_cols)
# hanya drop kolom yang memang ada di dataset
drop_cols = [c for c in drop_cols if c in df.columns]

print("\n fitur yang dipakai: ")
# print(df.columns.difference(drop_cols))
print(newdf.head())
```

Data preprocessing was conducted to ensure that the input variables used for model training were both valid and relevant. The variable Diagnosis was designated as the target attribute, as it represents the clinical outcome to be predicted by the machine-learning model. Several non-informative attributes such as PatientID, Ethnicity, and DoctorInCharge were excluded from the feature set because they were not functional and did not influence the model training. Prior to removal, the script verifies the presence of these attributes within the dataset to prevent runtime errors. By eliminating irrelevant identifiers, the resulting feature set consists solely of medically meaningful predictors, thereby improving the reliability and validity of the model's training process.

```
  scaler = StandardScaler()
  X_scaled = scaler.fit_transform(X)
```

The subsequent stage of preprocessing included standardization of the dataset to ensure all predictor variables were on an equal scale. This normalization procedure relied only on the statistical properties of the training set to avoid information leakage and ensured integrity of the evaluation process. This step is critical for Artificial Neural Networks (ANNs) because ANNs are known to be highly sensitive to scale differences among features. This can cause unstable training dynamics and poor convergence rates. However, after applying StandardScaler to the subsequent subsets of the data, the model trained and tested faster and more reliably with reduced dominance of each feature.

```
  X_train, X_test, y_train, y_test = train_test_split(
      X_scaled, y, test_size=0.2, random_state=42, stratify=y
  )
```

In the subsequent step of preprocessing, the data is split into training and testing samples. The data is split into 80% training data and 20% testing data. This is to ensure that there is enough data to learn while holding some data to test the model.

Stratified sampling is used to ensure that the data is equally distributed in both the training and testing samples. This is to avoid bias and to test the model using representative data.

4.1.2 Build and train random forest algorithm

```
grid_search = GridSearchCV(
    estimator=rf,
    param_grid=param_grid,
    cv=5,
    scoring='accuracy',
    n_jobs=-1,
    verbose=1
)
grid_search.fit(X_train, y_train)
```

The GridSearchCV approach is utilized during this phase to optimize the Random Forest model and discover the most suitable hyperparameters. Options 100, 200 and 500 for the number of trees in forest while max_depth is chosen using none, 10, and 20. The hyperparameters are analyzed using values that have been recognized as effective in previous studies, as reported in various journals.eu. A random number will be used to initiate the Random Forest model, which can be repeated. By cross-validating the training data with 5 times, the GridSearchCV approach can test each hyperparameter combination to ensure a stable model is chosen.

4.1.2 Build ANN

```
    ann_model = Sequential([
        Dense(128, input_dim=X_train.shape[1], activation='relu'),
        Dense(64, activation='relu'),
        Dense(64, activation='relu'),
        Dense(32, activation='relu'),
        Dense(16, activation='relu'),
        Dropout(0.3),
        Dense(1, activation='sigmoid')
    ])
```

An Artificial Neural Network (ANN) with multiple layers will be created to model the nonlinear trends in the Alzheimer's disease dataset. The network will be created with the maximum number of neurons in the initial layer to process the high-dimensional features of the input variables, which will be progressively reduced in the remaining layers. The inclusion of ReLU activation functions will ensure the smooth flow of the gradient during training of the network. To avoid the issue of over-specialization in the network, the dropout function will be introduced to avoid the overfitting of the network. To make the binary classification capable of predicting the probabilities, the final layer will include the sigmoid function.

4.1.3 Training ANN

```
    history = ann_model.fit(
        X_train, y_train,
        validation_data=(X_test, y_test),
        epochs=200,
        batch_size=64,
        class_weight=class_weights,
        callbacks=[early_stop],
        verbose=1
)
```

During the training process, the Artificial Neural Network is trained on the training data set through multiple epochs in order to enable it to learn the intricate patterns that exist in the data. A large enough number is defined in order to enable learning,

while early stopping is employed in order to check the validation performance. This prevents overfitting. In order to make the weight update stable, the mini-batch learning approach is employed. Class weighting is also employed in order to handle the issue of class imbalance, which involves giving more importance to the instances from the minority class during the learning process. The validation data is employed in order to check the generalization performance, which ensures that the final model is accurate.

4.2 Results

This subsection presents the classification results obtained from the Random Forest and Artificial Neural Network (ANN) models after all preprocessing and model training stages were completed. Model performance is evaluated using a confusion matrix to provide a detailed view of classification outcomes, including correct and incorrect predictions for both healthy individuals and Alzheimer's patients. The results are discussed by comparing each model's ability to correctly identify Alzheimer's cases and healthy samples, highlighting their respective strengths and limitations. This analysis serves as the basis for understanding the effectiveness of each model in supporting Alzheimer's disease classification.

The results from both models are summarized in the following table:

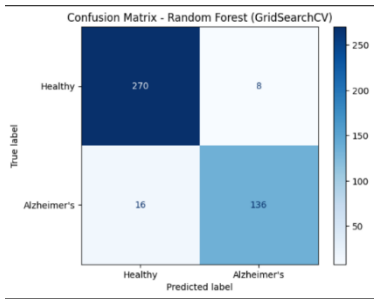

Figure 4. Confusion matrix - Random forest

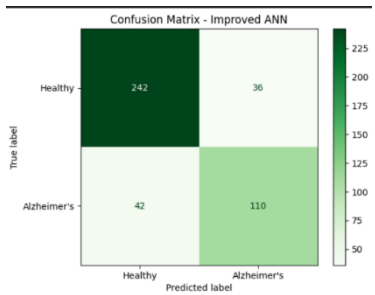

Figure 5. Improved AAN

Based on the confusion matrix, the Random Forest model shows a very high number of True Negatives, indicating strong performance in correctly identifying healthy individuals, but it still has limitations in detecting Alzheimer's cases due to the presence of False Negatives, where some affected patients are misclassified as healthy. In contrast, the Improved ANN produces a higher number of True Positives and fewer False Negatives, suggesting that it is more effective at identifying patients who actually have Alzheimer's disease, although this improvement comes with an increase in False Positives. Overall, the ANN demonstrates better sensitivity in detecting Alzheimer's cases, which is particularly important in medical classification tasks where missed diagnoses should be minimized.

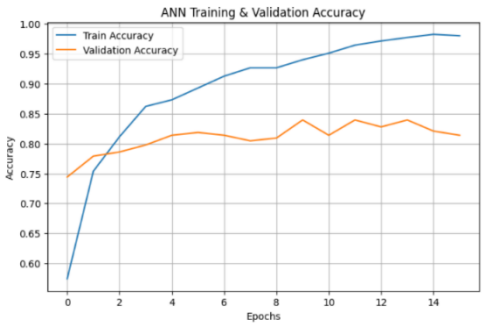

Figure 64. ANN training and validation accuracy

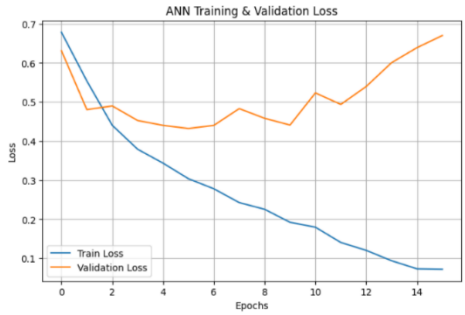

Figure 7. ANN training and validation loss

The first graph shows training and validation accuracy. While training accuracy steadily increases, validation accuracy improves initially but then fluctuates and slightly decreases, indicating overfitting. This suggests the model performs well on training data but struggles to generalize, and may benefit from techniques like early stopping, regularization, or more data.

The second graph shows training and validation loss. Training loss decreases consistently, but validation loss drops only early and then rises, further confirming overfitting. The model learns the training patterns well but performs worse on unseen data.

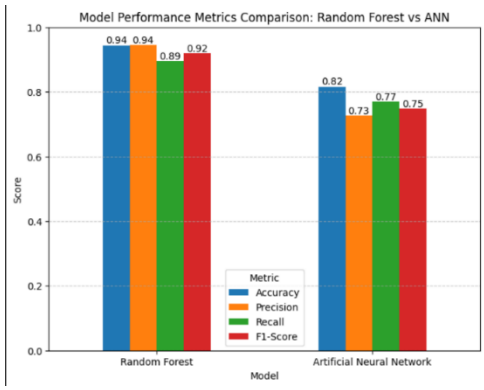

Figure8. Model performance metrics comparison: random forest vs ANN

This bar chart compares the performance metrics of the Random Forest and Artificial Neural Network models using four measures: accuracy, precision, recall, and F1-score. For the Random Forest model, the accuracy and precision both reach approximately 0.94, the recall is around 0.89, and the F1-score is about 0.92. For the Artificial Neural Network model, the accuracy is approximately 0.82, the precision is about 0.73, the recall is around 0.77, and the F1-score is close to 0.75. Each metric is represented by a different colored bar, and the values are displayed above the bars to show the exact scores for each model.

4.3 Discussion

The experimental results highlight how the structural differences between the models affect Alzheimer's disease detection. Random Forest achieves high overall accuracy, largely by correctly classifying non-Alzheimer's cases (high true negatives). Its ensemble decision boundaries favor dominant patterns in structured data, making it conservative for minority or borderline cases. Consequently, it produces many false negatives, limiting its usefulness for early detection where missing positive cases carries high risk.

In contrast, the ANN captures flexible, non-linear relationships in the data. Through iterative weight updates and class-weighted training, it identifies subtle feature combinations associated with Alzheimer's, resulting in more true positives and fewer false negatives. While overall accuracy is lower than Random Forest, ANN prioritizes sensitivity, making it better suited for early medical screening where detecting actual cases is critical.

5. CONCLUSION

The Random Forest (RF) and Artificial Neural Network (ANN) models show clear differences in accuracy, interpretability, and feasibility for early Alzheimer's diagnosis. Experimental results indicate that RF outperforms ANN in accuracy, precision, recall, and F1-score, demonstrating stronger overall classification performance. RF is also more interpretable, as its decision-making can be traced through feature importance and tree-based rules, which is valuable for clinical decision-making. In contrast, ANN acts as a black-box, making its internal processes harder to interpret, though it can model complex nonlinear relationships. RF is easier to train, requires less preprocessing, and is less sensitive to parameter tuning, making it more practical for real-world clinical use.

Both models contribute differently: RF provides reliable, stable predictions by aggregating multiple decision trees, while ANN can capture complex patterns in high-dimensional or nonlinear medical data. Despite lower performance in this study, ANN highlights the potential of deep learning when larger datasets and optimization are available.

Overall, RF emerges as the more effective and practical choice for clinical applications due to its predictive performance, interpretability, and ease of implementation. Model selection for early Alzheimer's diagnosis should consider not only accuracy but also interpretability and practical feasibility. Future research may expand datasets and integrate these models into clinical decision-support systems to improve usability and early detection outcomes.

REFERENCES

[1] Muhamad Yusuf Ismail, Andi Sunyoto, dan Agus Purwanto, 2024, Klasifikasi Penyakit Alzheimer pada Citra Medis Magnetic Resonance Images dengan Arsitektur Densenet121. Diakses dari

[2] Alda Amalia Mortara , Mitta Permatasari , Anita Desiani , Yuli Andriani , Muhammad Arhami, 2023, Perbandingan Algoritma C4.5 dan Adaptive Boosting dalam Klasifikasi Penyakit Alzheimer

[3] Maurino Putra, Erwin Harahap, 2024, Machine Learning pada Prediksi Kelulusan Mahasiswa Menggunakan Algoritma Random Forest

[4] Anjar Wanto, Sarjon Defit , Agus Perdana Windarto, 2021, Algoritma Fungsi Perlatihan pada Machine Learning berbasis ANN untuk Peramalan Fenomena Bencana

[5] Desmarita Leni , Helga yermadona , Ade Usra Berli , Ruzita Sumiati , Haris, 2022, Pemodelan Machine LearningUntuk Memprediksi Tensile Strength Aluminium Menggunakan Algoritma Artificial Neural Network(ANN)

[6] Euis Saraswati , Yuyun Umaidah, Apriade Voutama, 2021, Penerapan Algoritma Artificial Neural Network untuk Klasifikasi Opini Publik Terhadap Covid-19

[7] Ibnu Afdhal, Rahmad Kurniawan, Iwan Iskandar, Roni Salambue, Elvia Budianita, Fadhilah Syafria, 2022, Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia

[8] Muhamad Malik Mutoffar, Muchammad Naseer , Ariansyah Fadillah, 2022 KLASIFIKASI KUALITAS AIR SUMUR MENGGUNAKAN ALGORITMA RANDOM FOREST

[9] Muhamad Fajar Yudhistira Herjanto , Carudin , 2024, ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI SIREKAP PADA PLAY STORE MENGGUNAKAN ALGORITMA RANDOM FOREST CLASSIFER

[10] Ilham Kurniawan, Duwi Cahya Putri Buani, Abdussomad, Widya Apriliah, Rizal Amegia Saputra,2023IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK MENENTUKAN PENERIMA BANTUAN RASKIN