ITECX Journal International

e-ISSN: 3149-9147

Application of Machine Learning Algorithms in Early Detection of Alzheimer's Disease

Abdunabiyeva Diyora

National University of Uzbekistan named after Mirzo Ulugbek, Master's Student in Artificial Intelligence, abdunabiyevadiyora048@gmail.com

ABSTRACT

Published: 30.08.2026

This study proposes a multi-criteria evaluation approach for assessing machine learning models in the early detection of Alzheimer's disease. Conventional evaluation based on individual metrics such as accuracy, sensitivity, specificity, and ROC-AUC does not fully reflect overall diagnostic performance. To address this issue, an integral evaluation metric is introduced by combining multiple performance measures into a single indicator. The approach is applied to several models, including Random Forest, Support Vector Machine, and Convolutional Neural Network, using clinical and imaging data. The results show that the proposed method provides a more balanced and reliable assessment of model performance, enabling effective model selection. The approach can be utilized in medical decision-support systems to improve the accuracy and reliability of early diagnosis. Keywords: alzheimer's disease, machine learning, model evaluation, integrated metric, multi- criteria analysis

doi: 10.68209/hf4kh98

Keywords: *Alzheimer's disease, machine learning, model evaluation, integral indicator, multi-criteria analysis*

Introduction

Alzheimer's disease (AD) is the most common form of dementia among degenerative diseases of the central nervous system and is characterized by a gradual decline in cognitive functions, memory, and thinking abilities. According to reports by the World Health Organization, the increasing number of patients with dementia worldwide makes early detection and the development of effective diagnostic methods an important scientific problem [1].

Traditional diagnostic approaches, including neuropsychological tests, magnetic resonance imaging (MRI), and positron emission tomography (PET), often provide high accuracy at relatively late stages of the disease [2]. However, detecting subtle changes that occur in the initial stages remains challenging. For this reason, approaches based on artificial intelligence

and machine learning have become widely used in medical diagnostic systems in recent years [3].

In particular, machine learning algorithms make it possible to identify hidden patterns through the comprehensive analysis of large volumes of clinical, imaging, and genetic data. Studies show that models such as Random Forest, Support Vector Machine, and Convolutional Neural Network are highly effective in detecting Alzheimer's disease [4–6]. In particular, deep learning models based on MRI images are expanding the possibilities for detecting the disease at an early stage [5].

Therefore, there is a need to develop integrated approaches that comprehensively consider multiple criteria when evaluating machine learning models. The main objective of this study is to propose an approach that combines several diagnostic metrics into a single integral indicator for evaluating the effectiveness of machine learning models used in the early detection of Alzheimer's disease. This approach enables an objective comparison of different algorithms and helps identify the most balanced and clinically reliable model.

1. Research Methods

The main objective of this study is to make the evaluation of the effectiveness of machine learning models used in the early detection of Alzheimer's disease more objective and comprehensive. In recent years, approaches based on artificial intelligence have been widely applied in medical diagnostic systems, demonstrating strong results in detecting diseases at an early stage [1,2].

However, existing studies usually evaluate model performance using individual evaluation metrics—accuracy, sensitivity, specificity, and ROC-AUC [3]. This approach does not fully reflect a model's overall diagnostic ability because each metric represents only a particular aspect of the model. As a result, clinically unbalanced models that fail to adequately identify certain important cases may be selected [4].

For this reason, this study proposes an integrated approach based on multi-criteria evaluation. This approach enables a comprehensive comparison of different machine learning models, including Random Forest, Support Vector Machine, and Convolutional Neural Network, using a single criterion [5–7].

The proposed method is based on the concept of multi-criteria decision making (MCDM) and supports the selection of an optimal model by combining different evaluation criteria into a

single integral indicator [8]. The approach is also related to ensemble concepts, whereby combining several metrics increases the stability of the evaluation [9].

This approach can also be applied in multimodal systems capable of working with different types of data, including imaging, clinical, and cognitive data [10]. As a result, model selection does not depend on a single criterion but is directed toward comprehensive and clinically justified decision making.

2. Mathematical Formulation of the Proposed Method

To comprehensively evaluate model performance, a functional approach is proposed that combines several metrics into a single integral indicator:

where:

— accuracy,

— sensitivity,

— specificity,

— area under the ROC curve,

— corresponding weighting coefficients.

The weighting coefficients satisfy the following condition:

The main evaluation metrics are defined as follows:

where , , , respectively represent correct and incorrect classification results.

The proposed integral indicator makes it possible to evaluate model performance based on multiple criteria, helping identify the most balanced and clinically reliable model.

3. Database

The Alzheimer’s Disease Neuroimaging Initiative (ADNI) database was used in this study to address the problem of early detection of Alzheimer’s disease. This dataset is widely used in medical research and contains multimodal data [1].

The following types of data were used in the study:

— MRI images for analyzing brain structure

— Clinical and demographic indicators

— Cognitive test results, including MMSE and others

The subjects were divided into three main classes: cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD).

The data were divided into training and test sets at an 80/20 ratio, respectively, and k-fold cross-validation was applied to ensure model stability [2]. Balancing methods were used to reduce class imbalance.

An approach based on a Convolutional Neural Network was used to analyze imaging data, while tabular data were provided to classical machine learning models.

4. Data Preprocessing

Data preprocessing is the stage of preparing initial data before they are provided to a machine learning model. Its main purpose is to improve data quality, eliminate unnecessary or incomplete values, and ensure stable and accurate model operation. Preprocessing is particularly important in medical diagnostic problems because the quality of the initial data directly affects the reliability of model results.

In this study, missing values in tabular data were filled, features were normalized, and important features were selected. For MRI images, procedures were performed to standardize dimensions, reduce noise, and improve image quality. In addition, data augmentation methods were used to improve the model's generalization ability. These steps help increase the stability of model training and improve the accuracy of results.

5. Results

The effectiveness of different machine learning models was evaluated in the study. In particular, the performance of Random Forest, Support Vector Machine, and Convolutional Neural Network models was compared.

The main evaluation metrics—Accuracy, Sensitivity, Specificity, and ROC-AUC—were calculated for each model. The values of the proposed integrated evaluation indicator were also determined, and a comprehensive evaluation of the models was performed.

The following table presents model performance:

Model	Accuracy	Sensitivity	Specificity	AUC	Q
Random Forest	0.91	0.88	0.92	0.93	0.91
SVM	0.89	0.85	0.90	0.91	0.89
CNN	0.94	0.92	0.93	0.95	0.94

The results show that the CNN model achieved the highest results across all metrics. At the same time, the proposed indicator enabled a balanced evaluation of the models.

When individual metrics were considered separately, some models were observed to outperform others on only one indicator. For example, the Random Forest model provided high specificity, whereas the CNN model demonstrated an advantage in sensitivity and AUC.

The proposed integrated evaluation function made it possible to assess models in terms of overall effectiveness and identify the optimal model.

6. Discussion

The findings confirm the high effectiveness of machine learning algorithms in the early detection of Alzheimer's disease. In particular, the Convolutional Neural Network model achieved better results than the other models because of its ability to automatically extract features from MRI images.

At the same time, the Random Forest model also produced stable and reliable results and was observed to be particularly effective when working with clinical data. Although Support Vector Machine performed well on small datasets, it showed relatively lower effectiveness for complex imaging data.

The proposed integrated evaluation approach overcame the limitations of individual metrics and enabled a comprehensive and objective comparison of models. This approach is particularly important in medical diagnostic problems because incorrect decisions may lead to clinical errors.

The results show that relying on a single metric is insufficient when selecting a model; several criteria need to be considered together. From this perspective, the proposed method can be regarded as an effective tool for improving clinical decision-making systems.

Conclusion

This study examined the evaluation of the effectiveness of machine learning models in the early detection of Alzheimer's disease. The results showed that the Convolutional Neural Network model provides the highest accuracy using imaging data, while Random Forest produces stable results. The proposed integrated evaluation approach enables a comprehensive and balanced evaluation of models and helps select the optimal model. This approach is important for improving decision-making processes in medical diagnostic systems.

References

Baltrušaitis, T., Ahuja, C., Morency, L.P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 41*(2), 423–443. https://doi.org/10.1109/TPAMI.2018.2798607

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Dietterich, T.G. (2000). Ensemble methods in machine learning. Lecture Notes in Computer *Science, 1857*, 1–15. https://doi.org/10.1007/3-540-45014-9_1

Jack, C.R., Bennett, D.A., Blennow, K., et al. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia, 14(*4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018

LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature, 521,* 436–444. https://doi.org/10.1038/nature14539

Mueller, S.G., Weiner, M.W., Thal, L.J., et al. (2005). The Alzheimer's Disease Neuroimaging Initiative. *Neuroimaging Clinics of North America, 15*(4), 869–877. https://doi.org/10.1016/j.nic.2005.09.008

Noble, W.S. (2006). What is a support vector machine? *Nature Biotechnology, 24*, 1565–1567. https://doi.org/10.1038/nbt1206-1565

Powers, D.M.W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness and correlation. *Journal of Machine Learning Technologies, 2*(1), 37– 63.

Saito, T., Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE, 10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Snoek, J., Larochelle, H., & Adams, R. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, *25*.

Triantaphyllou, E. (2000). Multi-criteria decision making methods. In *Multi-criteria decision making methods: A comparative study* (pp. 5-21). Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-3157-6