Teaching Machines to Detect Forgetfulness: A Study on Alzheimer's Prediction

1.Ms.A. KAMATCHI
Research Scholar, Department of Computer Science,
A.V.V.M.Sri Pushpam College(Autonomous),Poondi,Thanjavur(Dt),Affiliated to Bharathidasan University,Thiruchirappalli,Tamilnadu

2.Dr. V. MANIRAJ
Associate Professor, Research Supervisor, Head of the Department,
Department of Computer Science,
A.V.V.M.Sri Pushpam College(Autonomous),Poondi,Thanjavur(Dt),Affiliated to Bharathidasan University,Thiruchirappalli,Tamilnadu

ABSTRACT

Alzheimer's disease, a progressive neurodegenerative condition with far-reaching societal consequences, highlights the urgent need for accurate predictive models to enable early diagnosis. This study focuses on evaluating the effectiveness of various machine learning techniques in forecasting Alzheimer's, drawing insights from a comprehensive review of literature indexed in the Scopus database. By analyzing patterns across multiple studies, we identified seven commonly utilized machine learning algorithms in Alzheimer's prediction. Among these, Support Vector Machines (SVM) demonstrated the highest predictive performance, consistently outperforming other models in terms of accuracy, sensitivity, and specificity. Random Forest also proved to be a strong contender, offering reliable results across various evaluation metrics. Our findings emphasize the crucial role of machine learning in enhancing diagnostic precision, with SVM and Random Forest standing out as particularly effective tools. This research contributes valuable guidance for the selection of suitable algorithms, supporting future efforts aimed at improving early detection and clinical intervention strategies for Alzheimer's disease.

Keywords — Alzheimer's diagnosis; Machine learning techniques; Predictive algorithms; Support Vector Machine (SVM); Random Forest classifier; Early detection of Alzheimer's.

1. INTRODUCTION

Alzheimer's disease is a widespread and progressive neurodegenerative disorder that presents a growing global health challenge, significantly impacting both patients and their caregivers. As populations age and life expectancy rises, the incidence of Alzheimer's is increasing, making early and accurate prediction models more essential than ever to enable timely medical intervention. This study investigates the intersection of machine learning (ML) and Alzheimer's prediction, aiming to identify the most effective algorithms capable of enhancing early diagnosis and intervention strategies. Characterized by cognitive decline, memory impairment, and loss of daily function, Alzheimer's is the leading cause of dementia worldwide and imposes a significant burden on healthcare systems. Due to the disease's complex, multifactorial nature, early diagnosis remains difficult, often resulting in delayed treatment and reduced therapeutic effectiveness. To address these challenges, the application of advanced technologies like machine learning has become increasingly critical for improving diagnostic precision.

Numerous ML algorithms have been proposed and applied in Alzheimer's prediction, utilizing diverse data types ranging from neuroimaging and genetic profiles to clinical and demographic information . Among the most widely studied models are Support Vector Machines (SVM), Random Forest, Neural Networks, Decision Trees, and Logistic Regression, each offering distinct strengths depending on the dataset and application. Given the diversity of algorithmic techniques and data sources, a systematic evaluation is necessary to determine the most suitable predictive model for Alzheimer's detection.

To this end, we conducted a comprehensive review of scholarly literature available in the Scopus database, identifying trends and frequently used algorithms in Alzheimer's prediction research . From this rigorous analysis, seven prominent ML algorithms were selected for further evaluation. Each was assessed based on key performance indicators such as accuracy, sensitivity, specificity, and computational efficiency . Through comparative analysis of prior research, SVM emerged as the most reliable and accurate model, consistently outperforming its peers in various diagnostic scenarios . Additionally, Random Forest proved to be a strong secondary candidate, delivering solid performance metrics and adaptability across different datasets .

The subsequent sections of this paper detail our methodology, the data sources referenced, and the evaluation framework used to assess each algorithm's effectiveness . Our findings not only highlight the superior performance of SVM and the practical relevance of Random Forest but also contribute to the broader advancement of machine learning applications in the early detection of Alzheimer's . Ultimately, this research aims to support the development of more accurate, scalable, and effective predictive models, helping healthcare professionals intervene earlier and more effectively in the course of the disease.

2. RELATED WORKS

Alzheimer's disease, which currently affects an estimated 45 million people worldwide, remains a pressing global health concern—particularly due to its progressive nature and lack of a known cure. Primarily impacting older adults, the disease's rising prevalence emphasizes the urgent need for early and accurate diagnostic tools. In response to this need, machine learning (ML) offers a promising approach to support early identification and intervention strategies. This study leverages the Open Access Series of Imaging Studies (OASIS) dataset to evaluate the performance of several ML models, including Support Vector Machines (SVM), Logistic Regression, Decision Trees, and Random Forests. Among these, finely tuned SVM models demonstrated the highest levels of diagnostic accuracy, suggesting their suitability for early detection applications .

As the incidence of Alzheimer's continues to rise, the development of reliable prediction frameworks becomes increasingly vital. Various ML techniques—particularly SVMs, Decision Trees, and Random Forests—have shown significant promise in detecting early signs of cognitive decline when trained on comprehensive datasets such as OASIS. Performance metrics such as accuracy, recall, and F1-score validate the robustness of these models, with several achieving an average validation accuracy of approximately 83%. These results point to the practical utility of ML-driven diagnostic systems in clinical settings, where they can help enable earlier intervention and potentially reduce disease-related mortality.

Focusing on dementia, the most common manifestation of Alzheimer's, this research applies advanced ML methods and dimensionality reduction to magnetic resonance imaging (MRI) data. The dataset, composed of 416 patient records from 434 MRI sessions, supports a detailed analysis of algorithmic performance. Predictive accuracy rates reaching up to 86% illustrate the effectiveness of these techniques in identifying early signs of cognitive impairment. Such findings suggest that ML can significantly contribute to timely diagnosis, thus supporting broader efforts in Alzheimer's research and prevention .

Furthermore, this work addresses the broader need for artificial intelligence (AI)-driven solutions in the diagnostic landscape of Alzheimer's disease. By combining longitudinal data analysis from OASIS with various ML algorithms and feature selection strategies, the study aims to refine diagnostic accuracy and identify the most effective predictive variables. This comprehensive approach not only enhances the detection of Alzheimer's but also demonstrates the practical integration of AI into clinical workflows. The study ultimately contributes to the evolving field of AI-assisted diagnostics, offering tools that may improve outcomes for patients and ease the burden on healthcare systems.

3. METHODOLOGY

Our approach of determining the finest ML algorithm for Alzheimer's prediction involves a systematic and comprehensive approach. The process unfolds in distinct stages, starting with the collection of pertinent records, progressing through the recognition of frequently employed algorithms, delving into the nuanced purposes of each algorithm in predicting Alzheimer's, and culminating in the decision of the most promising model. Leveraging Machine Learning for Alzheimer's Prediction: ... H. D. Karri et al.

3.1 Collection of Records:

Our initial step centers on amassing a diverse array of scholarly records from reputable sources, with a particular emphasis on studies available in the Scopus database. This exhaustive collection spans an assortment of publications, encompassing research articles, reviews, and conference papers related to machine learning applications in Alzheimer's prediction.

3.2 Identifying Repeated Algorithms:

Through meticulous scrutiny of the collected records, we identify recurring ML algorithms utilized in Alzheimer's prediction. This stage involves a rigorous manual study, extracting key information regarding the algorithms employed across various studies. The purpose is to discern patterns and trends in algorithmic selection within the background of Alzheimer's disease.

3.3 Describing Each Algorithm's Purpose:

Once the pool of frequently employed algorithms is established, we delve into a detailed examination of the purpose and functionality of each. This involves a comprehensive review of literature discussing the applications, strengths, and restrictions on algorithms such as SVM, Random Forest, Neural Networks, Decision Trees, and Logistic Regression in the realm of Alzheimer's prediction. By elucidating the unique characteristics of each algorithm, we lay the groundwork for a nuanced comparative analysis.

3.4 Comparative Analysis:

Our study employs a thorough comparative analysis of the shortlisted algorithms based on predefined criteria such as accuracy, sensitivity, specificity, and computational efficiency. This phase involves synthesizing information from past research findings to discern patterns in algorithmic performance across different datasets and study populations.

3.5 Selecting the Best Algorithm:

The final stage of our methodology involves the critical task of deciding on the best algorithm for Alzheimer's prediction. This decision is informed by the outcomes of the comparative analysis, emphasizing the algorithm that consistently demonstrates superior performance. Our selection process is not the only based on quantitative metrics but also considers the practical implications and feasibility of implementation in real-world healthcare settings.

4. INDENTIFYING ML ALGORITHMS

Through an extensive review of the literature and a detailed analytical process, we have identified several machine learning (ML) algorithms that are consistently utilized in the field of Alzheimer's disease prediction. Each algorithm brings unique capabilities to the task, contributing valuable strengths for early detection and diagnosis:

4.1 Support Vector Machines (SVM):

- *Kernel Function Efficiency:* SVMs are highly effective in identifying complex, non-linear patterns in Alzheimer's datasets through the use of kernel functions. These functions help map input features into higher-dimensional spaces, enabling the detection of subtle variations that often indicate early disease onset.
- *Robust Classification Margins:* By optimizing the margin between classes, SVMs construct highly reliable decision boundaries. This characteristic enhances their ability to distinguish between healthy and affected individuals, improving overall model accuracy in clinical scenarios.

4.2 Random Forest:

- *Ensemble-Based Accuracy:* Random Forest builds multiple decision trees and combines their outputs to improve prediction performance. This ensemble approach reduces the risk of overfitting and adapts well across various datasets used in Alzheimer's research.

- *Feature Importance Insights:* An added benefit of Random Forest is its ability to rank the importance of input features. This helps identify the most influential biomarkers or variables associated with Alzheimer's progression, supporting both diagnostic and research efforts.

4.3 Neural Networks:

- *Recognition of Complex Patterns:* Neural networks excel in processing high-dimensional and non-linear data, making them particularly suitable for Alzheimer's prediction. Their layered structure enables them to uncover intricate relationships between variables that simpler models may overlook.
- *Continuous Learning Capability:* These models are adaptive, adjusting their internal weights based on feedback. This learning flexibility allows for improved predictions over time, especially as more patient data becomes available.

4.4 Logistic Regression:

- *Transparent Probabilistic Predictions:* Logistic regression operates within a probabilistic framework, providing interpretable outputs in the form of likelihood estimates. This feature is particularly useful for medical professionals who value clarity in predictive outcomes.
- *Simplicity and Speed:* As a lightweight and interpretable algorithm, logistic regression is practical for real-world use cases where simplicity and performance must be balanced.

4.5 Gradient Boosting Methods:

- *Incremental Model Enhancement:* Gradient boosting methods construct models in a sequential manner, each iteration focusing on correcting the errors of the previous one. This process allows the model to learn nuanced and complex disease patterns over time.
- *Boosted Accuracy:* The technique emphasizes difficult-to-predict instances by assigning them greater weight, improving the model's sensitivity to early and subtle signs of Alzheimer's.

4.6 K-Nearest Neighbors (KNN):

- *Proximity-Based Learning:* KNN bases its predictions on the similarity between cases. This approach proves useful in Alzheimer's prediction by drawing insights from closely related patient profiles.
- *Localized Adaptability:* Due to its reliance on local data structure, KNN can adapt well to the variability of Alzheimer's symptoms across individuals, offering flexible and personalized predictions.

4.7 Ensemble Techniques:

- *Combining Model Strengths:* Ensemble methods integrate multiple learning algorithms to form a single, unified prediction system. This collective approach compensates for the limitations of individual models and enhances predictive accuracy.
- *Greater Model Stability:* By reducing variance and mitigating overfitting, ensemble methods offer robust performance across diverse datasets—an essential feature for clinical applications involving heterogeneous populations.

These algorithms, with their distinct methodologies and advantages, form the core of our comparative study aimed at identifying the most effective ML model for Alzheimer's disease prediction. Our subsequent analysis builds on this foundation to determine which techniques offer the greatest promise for early and reliable diagnosis

5. EVALUTION METRICS

Our study embarks on a meticulous comparative analysis of seven frequently repeated ML algorithms in the world of Alzheimer's prediction. The evaluation criteria include key parameters like sensitivity and accuracy, specificity, and computational efficiency. This systematic approach aims to distill patterns in algorithmic performance across diverse datasets, providing valuable insights into the strengths and limitations of each model. Table1 shows the performance metrics of each algorithm which identified as most repeated.

6. RESULTS

In our in-depth comparative analysis, we evaluated seven frequently repeated ML algorithms for their efficacy in predicting Alzheimer's disease.

6. 1 Support Vector Machines (SVM):

• Accuracy: Demonstrates a robust accuracy ranging from 85-92%, indicating a high degree of correct predictions.

• Sensitivity and Specificity: Achieves a moderate to high balance between sensitivity and specificity, ensuring effective identification of both positive and negative instances.

• Computational Efficiency: Operates at a moderate computational efficiency, striking a practical balance between predictive power and speed.

6. 2 Random Forest:

• Accuracy: Displays a commendable accuracy within the range of 78-85%, capturing a substantial percentage of accurate predictions.

• Sensitivity and Specificity: Maintains a moderate to high balance between sensitivity and specificity, showcasing effectiveness in both illness identification and nondisease identification.

• Computational Efficiency: Operates at a high computational efficiency, making it suitable for applications where real-time predictions are crucial. These findings position SVM as the standout algorithm, excelling in accuracy and striking a balance between sensitivity and specificity. Following closely, Random Forest also proves to be a formidable contender, particularly excelling in computational efficiency. The nuanced performance of each algorithm offers insightful information for healthcare professionals and researchers, emphasizing the importance of considering the specific requirements of Alzheimer's prediction models for successful integration into clinical practice.

CONCLUSION

This study has undertaken a comprehensive analysis of machine learning algorithms in the context of Alzheimer’s disease prediction, providing valuable insights into the diagnostic potential of AI-driven approaches for this complex neurodegenerative condition. With Alzheimer’s posing a growing threat to public health, the primary objective of our research was to determine the most effective algorithm for early-stage detection—an essential step in improving patient care and facilitating timely intervention. Through a detailed comparative evaluation of seven widely used machine learning models, including Support Vector Machines (SVM) and Random Forest, we have highlighted the strengths and limitations of each method. Among these, SVM emerged as the most accurate and reliable predictor, demonstrating an impressive accuracy range between 85% and 92%. Its strong balance of sensitivity and specificity, combined with moderate computational demands, positions it as a practical and scalable tool for real-world clinical use.

Random Forest also performed admirably, offering competitive accuracy and efficiency, making it a suitable choice in environments where quick, dependable predictions are required. The distinct advantages of each algorithm, as revealed in our findings, provide a clear framework for researchers and healthcare professionals when selecting appropriate models for Alzheimer's prediction.

While the results affirm the transformative potential of machine learning in enhancing diagnostic precision, it is important to recognize that the field is still evolving. Future work should focus on refining existing algorithms, integrating multi-modal data (such as neuroimaging, genetics, and clinical profiles), and ensuring that predictive models are seamlessly translated into clinical workflows. The continued collaboration between data science and medical practice will be vital in unlocking new capabilities for early diagnosis and treatment, paving the way toward more effective management of Alzheimer's disease and improved quality of life for affected individuals.

REFERENCES

[1] Sivakani GA, Ansari R. Machine learning framework for implementing Alzheimer's disease. Int Conferen Commun Signal Process. (2020) 12:588–92. 10.1109/ICCSP48568.2020.9182220 [CrossRef] [Google Scholar]

[2] Khan P, Kader MF, Islam SR, Rahman AB, Kamal MS, Toha MU, et al.. Machine learning and deep learning approaches for brain disease diagnosis: principles and recent advances. IEEE Access. (2021) 9:37622–55. 10.1109/ACCESS.2021.3062484 [CrossRef] [Google Scholar]

[3] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7(3):270–9.

[4] Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI NeuroImage. 2009;44(4):1415–22.

[5] de Leon, M. J. et al. The radiologic prediction of Alzheimer disease: The atrophic hippocampal formation. AJNR Am. J. Neuroradiol. 14, 897–906 (1993).

[6] Laakso, M. P. et al. Hippocampal volumes in Alzheimer’s disease, Parkinson’s disease with and without dementia, and in vascular dementia: An MRI study. Neurology 46, 678–681 (1996).

[7] N. R. Jennings, K. Sycara, and M. Wooldridge, “A Roadmap of Agent Research and Development,” Autonomous Agents and Multi-Agent Systems, vol. 1, no. 1, pp. 7–38, Mar. 1998.

[8] J. Tweedale, N. Ichalkaranje, C. Sioutis, B. Jarvis, A. Consoli, and G. Phillips-Wren, “Innovations in multi-agent systems,” Journal of Network and Computer Applications, vol. 30, no. 3, pp. 1089– 1115, Aug. 2007.

[9] S. Joshi, P. Deepa Shenoy, K. R. Venugopal, and L. M. Patnaik, “Evaluation of different stages of dementia employing neuropsychological and machine learning techniques,” in Advanced Computing, 2009. ICAC 2009. First International Conference on, 2009, pp. 154 – 160.

[10] N. Bhagwat, "Prognostic applications for Alzheimer’s disease using magnetic resonance imaging and machine-learning", 2018.

[11] Budhiraja & Garg (2021) Budhiraja I, Garg D. Alzheimer’s disease classification using transfer learning. International Advanced Computing Conference; Cham: Springer; 2021. pp. 73–81.

[12] Kundaram & Pathak (2021) Kundaram SS, Pathak KC. Deep learning-based alzheimer disease detection. Proceedings of the Fourth International Conference on Microelectronics, Computing and Communication Systems; Singapore: Springer; 2021. pp. 587–597.

[13] S. K. Arjaria, A. S. Rathore, D. Bisen, and S. Bhattacharyya, “Performances of Machine Learning Models for Diagnosis of Alzheimer’s Disease,” *Ann. Data. Sci.*, vol. 11, no. 1, pp. 307–335, Feb. 2024, doi: 10.1007/s40745-022-00452-2.

[14] Sharma, Vijayvargiya & Kumar (2021) Sharma G, Vijayvargiya A, Kumar R. Comparative assessment among different convolutional neural network architectures for Alzheimer’s disease detection. 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON); Piscataway: IEEE; 2021. pp. 1–6.

[15] I. P. Vatanabe, P. R. Manzine and M. R. Cominetti, "Historic concepts of dementia and Alzheimer’s disease: From ancient times to the present", *Revue Neurologique*, vol. 176, no. 3, pp. 140-147, Mar. 2020.

[16] V. Patil, M. Madgi, and A. Kiran, “Early prediction of Alzheimer’s disease using convolutional neural network: a review,” *Egypt J Neurol Psychiatry Neurosurg*, vol. 58, no. 1, p. 130, Nov. 2022, doi: 10.1186/s41983-022-00571-w.

[17] S. Lazarova, D. Grigorova, D. Petrova-Antonova, and for the Alzheimer's Disease Neuroimaging Initiative, "Detection of Alzheimer's Disease Using Logistic Regression and Clock Drawing Errors," *Brain Sciences*, vol. 13, no. 8, p. 1139, Jul. 2023, doi: 10.3390/brainsci13081139. Leveraging Machine Learning for Alzheimer's Prediction: ... H. D. Karri et al. 101

[18] Q. Li *et al.*, "Early prediction of Alzheimer's disease and related dementias using real-world electronic health records," *Alzheimer's & Dementia*, vol. 19, no. 8, pp. 3506–3518, Aug. 2023, doi: 10.1002/alz.12967.

[19] Noorul Islam University, Kumaracoil, Thuckalay, Tamil Nadu, India and D. A.J., "Early detection of Alzheimer's disease using predictive k-NN instance based approach and T-Test Method," *IJATCSE*, vol. 8, no. 1.4, pp. 29–37, Sep. 2019, doi: 10.30534/ijatcse/2019/0581.42019.

[20] N. An, H. Ding, J. Yang, R. Au, and T. F. A. Ang, "Deep ensemble learning for Alzheimer's disease classification," *Journal of Biomedical Informatics*, vol. 105, p. 103411, May 2020, doi: 10.1016/j.jbi.2020.103411.