Jurnal Informatika Ekonomi Bisnis

http://www.infeb.org

2024 Vol. 6 Iss. 4 Hal: 810-814 e-ISSN: 2714-8491

Advancing Alzheimer's Diagnosis: A Comparative Analysis of Deep Learning Architectures on Multidimensional Health Data

Gregorius Airlangga[1]✉

[1]Atma Jaya Catholic University of Indonesia

gregorius.airlangga@atmajaya.ac.id

Abstract

Alzheimer's Disease (AD) is a leading cause of disability among the elderly, with its prevalence projected to triple by 2050. Early detection remains critical for effective disease management, yet traditional diagnostic methods are often time-intensive and subjective. This study investigates the effectiveness of three machine learning architectures: Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) in detecting Alzheimer's Disease using a multidimensional dataset comprising demographic, lifestyle, medical, cognitive, and functional data from 2,149 patients. Each model was evaluated using 10-fold cross-validation, with performance metrics including accuracy, precision, recall, and F1-score. The CNN model demonstrated superior performance, achieving an average accuracy of 88.65%, surpassing both the MLP (84.41%) and LSTM (75.57%) models. These results highlight CNNs' capability to effectively extract spatial patterns in health data, making them a promising tool for Alzheimer's diagnosis. In contrast, LSTM underperformed due to the lack of temporal relationships in the dataset. This study underscores the importance of aligning model architecture with dataset characteristics and provides a foundation for integrating machine learning into clinical workflows. Future work will focus on hybrid architectures and real-world validation to enhance diagnostic accuracy and scalability.

Keywords: Alzheimer's Disease, Deep Learning, Machine Learning, Diagnostic Tools, Multidimensional Health Data.

1. Introduction

Alzheimer's Disease (AD) remains a significant global health challenge due to its progressive nature and the devastating impact it has on cognitive and functional abilities [1]. It is a major cause of disability among elderly individuals, with a growing prevalence as the global population ages [2]. According to recent studies, the number of individuals affected by AD is projected to triple by 2050, emphasizing the need for better diagnostic and preventive strategies [3]. Traditional methods of diagnosing Alzheimer's Disease heavily rely on clinical observations and neuropsychological testing, which, while effective, are time-consuming, subjective, and often fail to detect early symptoms [4], [5], [6]. This underscores the need for innovative, data-driven approaches that leverage advances in machine learning to address these limitations [7]. By incorporating demographic, lifestyle, medical history, and cognitive data, machine learning models can offer a comprehensive analysis of risk factors and symptoms associated with AD, providing a robust framework for early detection [8].

Despite significant advancements in medical diagnostics, existing literature reveals critical gaps in the application of machine learning models for Alzheimer's detection. For instance, many prior studies have focused solely on individual predictors, such as genetic predisposition or cognitive tests, without integrating multiple dimensions of patient data. A study demonstrated the potential of logistic regression models using demographic and medical history data for AD diagnosis but reported moderate accuracy due to the lack of cognitive and functional assessment variables [9]. Similarly, other study applied neural networks to imaging data, achieving high precision but at the cost of computational resources, making the approach impractical for routine screenings [10]. Furthermore, certain research highlighted the need for addressing imbalanced datasets, as Alzheimer's cases are often underrepresented in broader health data, leading to skewed predictions [11]. These gaps reveal an opportunity to explore hybrid and ensemble models that can efficiently integrate diverse data types to improve diagnostic performance.

The state of the art in Alzheimer's research increasingly points towards deep learning models, which have shown promise in extracting complex patterns from multidimensional data. Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) models, for instance, have been successfully utilized in other domains of medical diagnostics, such as cancer detection and diabetes prediction [12], [13], [14]. However, their application in Alzheimer's detection has been limited by challenges such as overfitting, lack of interpretability, and insufficient validation on diverse datasets [15], [16], [17]. Moreover, while multi-layer perceptrons (MLPs) have been explored for classifying cognitive impairment, their performance often suffers due to the heterogeneity of Alzheimer's data [18]. The present research addresses these challenges by proposing a comprehensive evaluation of MLPs, CNNs,

Accepted: 21-11-2024 | Revision: 06-12-2024 | Publication: 31-12-2024 | doi: 10.37034/infeb.v6i4.1046

and LSTMs, combined with robust cross-validation techniques to ensure generalizability and reliability [19].

The urgency of advancing diagnostic tools for Alzheimer's Disease cannot be overstated. As the disease progresses, the cost of care escalates significantly, imposing a substantial economic burden on families and healthcare systems [20]. Early detection, enabled by predictive models, has the potential to mitigate these costs by facilitating timely interventions and better management of the disease [21]. Additionally, integrating machine learning into clinical workflows aligns with global initiatives to digitize healthcare and improve access to personalized medicine [22]. The ability to analyze complex interactions among demographic, lifestyle, medical, and cognitive factors is critical for identifying at-risk individuals and tailoring preventive measures accordingly [23], [24].

The primary goal of our research is to develop and evaluate machine learning models that leverage the rich diversity of data available in the Alzheimer's Disease Dataset to enhance diagnostic accuracy. Our approach integrates MLP, CNN, and LSTM architectures, each suited to different aspects of the dataset's multidimensional nature. By employing ten-fold cross-validation, we aim to ensure robust model evaluation and address the common issue of overfitting in Alzheimer's research. Our contribution lies in not only identifying the best-performing model but also providing insights into the interpretability and practical applicability of these models in real-world scenarios. This study sets itself apart by utilizing an extensive dataset that includes demographic, medical, cognitive, and functional assessments, making the findings highly relevant to clinical practice.

The remainder of this article is structured as follows. First, we present the materials and methods, detailing the dataset preparation, preprocessing techniques, and the architecture of the machine learning models used. Next, the results section provides a comparative analysis of the model performances, followed by an in-depth discussion of the implications, limitations, and potential future directions. Finally, we conclude by summarizing our key findings and emphasizing the significance of integrating machine learning into Alzheimer's diagnostics. Through this research, we aim to contribute to the growing body of knowledge in digital health, providing a scalable and efficient framework for addressing one of the most pressing challenges in modern healthcare.

2. Research Method

This section details the research methodology, elaborating on the dataset preparation, preprocessing techniques, and the architecture of the machine learning models. The dataset used in this study is the Alzheimer's Disease Dataset and can be downloaded from certain source, which contains detailed records of 2,149 patients, each identified by a unique Patient ID ranging from 4751 to 6900 [25]. Each record includes demographic, lifestyle, medical history, clinical, cognitive, and functional assessments, along with the diagnosis of Alzheimer's Disease. Formally, let the dataset be represented as $(D = \{(X_i, y_i)\}_{i=1}^{n})$, where $(X_i \in R^d)$ is the feature vector for the (i)-th patient, $(y_i \in \{0,1\})$ is the binary diagnosis label indicating the absence or presence of Alzheimer's Disease, and $(n = 2149)$ is the total number of samples.

The dataset preprocessing involves three main steps: removal of non-predictive features, normalization of the input features, and encoding of the target variable. First, features that do not contribute to the predictive task, such as PatientID and DoctorInCharge, were excluded, reducing the feature dimension to $(d = 28)$. The cleaned dataset is thus represented as $(D_{\text{cleaned}} = \{(X_i', y_i)\}_{i=1}^{n})$, where $(X_i' \in R^{28})$.

The second step involves normalization to standardize the feature values. Let (X_{ij}') represent the (j)-th feature of the (i)-th sample. Each feature (j) is transformed as presented in Equation 1.

$$X_{ij}'' = \frac{X_{ij}' - \mu_j}{\sigma_j} \tag{1}$$

where $(\mu_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij}')$ is the mean of feature (j), and $(\sigma_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_{ij}' - \mu_j\right)^2})$ is the standard deviation. This transformation ensures that all features have a mean of 0 and a standard deviation of 1, facilitating efficient optimization during model training. The target variable $(y \in \{0, 1\})$ is encoded using one-hot encoding, resulting in $(Y \in R^{n\times 2})$, where each row $(Y_i = [y_{i0}, y_{i1}])$ satisfies $(y_{i0} + y_{i1} = 1)$, with (y_{i0}) indicating absence $((y_i = 0)) and (y_{i1})$ indicating presence $((y_i = 1))$ of Alzheimer's Disease.

To ensure reliable model evaluation, the dataset was partitioned using (k)-fold cross-validation with $(k = 10)$. In each fold $(k \in \{1, \dots, 10\})$, the dataset is split into training $((X_{\text{train}}^k, Y_{\text{train}}^k))$ and testing $((X_{\text{test}}^k, Y_{\text{test}}^k))$ subsets, maintaining the class distribution. Let (n_k^{train}) and (n_k^{test}) denote the sizes of the training and testing subsets, respectively, such that $(n_k^{\text{train}} + n_k^{\text{test}} = n)$. Three machine learning architectures were employed: Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) network. Each architecture is tailored to different aspects of the dataset's multidimensional structure. The MLP architecture consists of three fully connected layers. The layer has $(d = 28)$ neurons corresponding to the feature dimension. The first hidden layer has $(h_1 = 128)$ neurons with ReLU activation, defined as $(f(x) = \max(0, x))$. Dropout regularization with a rate $(r = 0.3)$ is applied to reduce overfitting.

The second hidden layer contains ($h_2 = 64$) neurons, also with ReLU activation and dropout. The output layer has two neurons, corresponding to the two classes, and uses the softmax activation function as presented in Equation 2.

$$\sigma(z_i) = \frac{\exp(z_i)}{\sum_{j=1}^{2} \exp(z_j)} \quad (2)$$

where (z_i) is the (i)-th output logit. The CNN model begins with reshaping the input feature matrix ($X \in R^{n\times d}$) to ($X' \in R^{n\times d\times 1}$). The first convolutional layer applies ($f = 64$) filters of size ($k = 3$), followed by a second convolutional layer with ($f = 32$) filters of size ($k = 3$). Mathematically, the convolution operation is given by $z_{i,j} = \sum_{p=0}^{k-1} w_p x_{i,j+p} + b$, where ($z_{i,j}$) is the output of the ($j$)-th convolution for the (i)-th sample, (w_p) are the kernel weights, and (b) is the bias term. The outputs are flattened into a vector of size (v) and passed through a dense layer with 64 neurons before the softmax output layer.

The LSTM model captures sequential dependencies in the input features. The input is reshaped as ($X' \in R^{n\times d\times 1}$). The first LSTM layer has ($u = 128$) units and processes input sequentially, updating the hidden state (h_t) and cell state (c_t) at each time step (t) as presented in Equation 3 and 4.

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f), \quad i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (3)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o), \quad c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (4)$$

where (f_t), (i_t), and (o_t) are the forget, input, and output gates, respectively, and ($\odot$) denotes element-wise multiplication. A second LSTM layer with ($u = 64$) units refines the representations. Dropout regularization with ($r = 0.3$) is applied after each LSTM layer, and the final output is passed through a dense layer with softmax activation. The performance of each model is evaluated using metrics such as accuracy, precision, recall, and F1-score. Let (TP), (FP), (TN), and (FN) denote the true positives, false positives, true negatives, and false negatives, respectively. The metrics are defined as presented in Equation 5, 6 ,7 and 8.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

This methodology ensures comprehensive evaluation and highlights the effectiveness of advanced machine learning architectures in detecting Alzheimer's Disease.

3. Result and Discussion

The results of this study demonstrate the effectiveness of different machine learning architectures: Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) in detecting Alzheimer's Disease as presented in the table 1. Each model was evaluated using four key performance metrics: accuracy, precision, recall, and F1-score. The metrics provide a comprehensive understanding of each model's predictive capability and generalization performance.

Table 1. Deep Learning's Performance Results

Model	Accuracy	Precision	Recall	F1 Score
MLP	0.8441	0.8430	0.8441	0.8417
CNN	0.8865	0.8884	0.8865	0.8862
LSTM	0.7557	0.7567	0.7557	0.7528

The MLP model achieved an average accuracy of 84.41%, a precision of 84.30%, a recall of 84.41%, and an F1-score of 84.17%. These results indicate that the MLP model performs consistently across all metrics, showing balanced precision and recall. This suggests that the MLP model effectively captures the linear and nonlinear relationships within the features, leveraging its fully connected architecture. However, while the MLP achieves reasonable performance, its metrics lag behind those of the CNN model, highlighting the limitations of simple dense architectures when handling complex, multidimensional data.

The CNN model outperformed the MLP in all metrics, achieving an average accuracy of 88.65%, precision of 88.84%, recall of 88.65%, and F1-score of 88.62%. The CNN's superior performance can be attributed to its ability to capture spatial patterns within the data through convolutional layers. By applying filters, the CNN extracts high-level feature representations that are likely more informative for the classification task. Additionally, the combination of convolutional layers and dense layers enables the CNN to generalize well across folds, as evidenced by its consistently high recall and F1-score. This makes the CNN model the most robust and accurate among the three architectures evaluated in this study.

In contrast, the LSTM model achieved an average accuracy of 75.57%, precision of 75.67%, recall of 75.57%, and F1-score of 75.28%. While LSTMs are particularly effective for sequential and temporal data, their performance in this study indicates limitations when applied to tabular, non-sequential datasets like the Alzheimer’s Disease Dataset. The relatively lower performance metrics suggest that the LSTM model struggled to extract meaningful dependencies or patterns, likely due to the absence of inherent temporal relationships in the data. Furthermore, the higher computational complexity of LSTMs compared to CNNs and MLPs may have contributed to less optimal performance during the cross-validation process. The comparison of these models highlights important insights into their suitability for Alzheimer’s Disease detection. The CNN model’s ability to outperform both MLP and LSTM architectures underscores the importance of leveraging spatial feature extraction techniques in high-dimensional datasets. In contrast, the lower performance of the LSTM emphasizes that model selection should align with the nature of the data; while LSTMs excel in time-series or sequential tasks, they may not be well-suited for tasks involving tabular data without temporal dimensions. These findings also emphasize the critical role of architecture design in optimizing predictive performance. For instance, the use of convolutional layers in the CNN allowed it to exploit feature hierarchies, while the dense layers in the MLP captured generalized patterns. On the other hand, the LSTM’s reliance on sequential processing appeared less effective given the structure of the input data. This suggests that future studies could explore hybrid models that combine CNNs with other architectures, such as MLPs, to further enhance predictive capabilities.

4. Conclusion

This study investigates the effectiveness of three machine learning architectures: Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) for detecting Alzheimer’s Disease using a comprehensive dataset that integrates demographic, lifestyle, medical, and cognitive features. Employing 10-fold cross-validation, the evaluation revealed that the CNN model outperformed the others, achieving the highest metrics (accuracy: 88.65%, F1-score: 88.62%), underscoring its robustness in capturing complex spatial patterns. The MLP showed reasonable performance (accuracy: 84.41%, F1-score: 84.17%), while the LSTM, more suited for temporal data, struggled with the dataset’s tabular nature, achieving lower metrics (accuracy: 75.57%, F1-score: 75.28%). These findings emphasize the importance of aligning model architecture with dataset characteristics, with CNNs proving particularly effective for complex feature interactions. The study contributes to data-driven healthcare by demonstrating the potential of machine learning models as diagnostic tools, with future directions including hybrid models, interpretability techniques, and validation on larger datasets to enhance clinical applicability and improve early detection and management of Alzheimer’s Disease.

References

[1] Abubakar, M. B., Sanusi, K. O., Ugusman, A., Mohamed, W., Kamal, H., Ibrahim, N. H., Khoo, C. S., & Kumar, J. (2022). Alzheimer’s disease: An update and insights into pathophysiology. *Frontiers in Aging Neuroscience, 14*, 742408. https://doi.org/10.3389/fnagi.2022.742408

[2] Dong, Y., Peng, R., Kang, H., Song, K., Guo, Q., Zhao, H., Zhu, M., Zhang, Y., Guan, H., & Li, F. (2022). Global incidence, prevalence, and disability of vertebral fractures: A systematic analysis of the global burden of disease study 2019. *The Spine Journal, 22*(5), 857–868. https://doi.org/10.1016/j.spinee.2021.12.007

[3] Nandi, A., Counts, N., Chen, S., Seligman, B., Tortorice, D., Vigo, D., & Bloom, D. E. (2022). Global and regional projections of the economic burden of Alzheimer's disease and related dementias from 2019 to 2050: A value of statistical life approach. *EClinicalMedicine, 51*, 101580. https://doi.org/10.1016/j.eclinm.2022.101580

[4] Dara, O. A., Lopez-Guede, J. M., Raheem, H. I., Rahebi, J., Zulueta, E., & Fernandez-Gamiz, U. (2023). Alzheimer’s disease diagnosis using machine learning: A survey. Applied Sciences, 13(14), 8298. https://doi.org/10.3390/app13148298

[5] Deng, Y., Wang, H., Gu, K., & Song, P. (2023). Alzheimer's disease with frailty: Prevalence, screening, assessment, intervention strategies and challenges. *Bioscience Trends, 17*(4), 283–292. https://doi.org/10.5582/bst.2023.01112

[6] Mahaman, Y. A. R., Embaye, K. S., Huang, F., Li, L., Zhu, F., Wang, J.-Z., Liu, R., Feng, J., & Wang, X. (2022). Biomarkers used in Alzheimer’s disease diagnosis, treatment, and prevention. *Ageing Research Reviews, 74*, 101544. https://doi.org/10.1016/j.arr.2021.101544

[7] Strielkowski, W., Vlasov, A., Selivanov, K., Muraviev, K., & Shakhnov, V. (2023). Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review. *Energies, 16*(10), 4025. https://doi.org/10.3390/en16104025

[8] Liss, J. L., Seleri Assunção, S., Cummings, J., Atri, A., Geldmacher, D. S., Candela, S. F., Devanand, D. P., Fillit, H. M., Susman, J., & Mintzer, J. (2021). Practical recommendations for timely, accurate diagnosis of symptomatic Alzheimer’s disease (MCI and dementia) in primary care: A review and synthesis. *Journal of Internal Medicine, 290*(2), 310–334. https://doi.org/10.1111/joim.13254

[9] González, D. A., Gonzales, M. M., Jennette, K. J., Soble, J. R., & Fongang, B. (2021). Cognitive screening with functional assessment improves diagnostic accuracy and attenuates bias. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 13*(1), e12250. https://doi.org/10.1002/dad2.12250

[10] Gangwal, A., Ansari, A., Ahmad, I., Azad, A. K., & Sulaiman, W. M. A. W. (2024). Current strategies to address data scarcity in artificial intelligence-based drug discovery: A comprehensive review. *Computers in Biology and Medicine, 179*, 108734. https://doi.org/10.1016/j.compbiomed.2023.108734

[11] Chang, C. Y., Slowiejko, D., & Win, N. (2024). Prediction and clustering of Alzheimer’s disease by race and sex: A multi-head deep-learning approach to analyze irregular and heterogeneous data. *Scientific Reports, 14*(1), 26668. https://doi.org/10.1038/s41598-024-26668-0

[12] Madan, P., Singh, V., Chaudhari, V., Albagory, Y., Dumka, A., Singh, R., Gehlot, A., Rashid, M., Alshamrani, S. S., & AlGhamdi, A. S. (2022). An optimization-based diabetes prediction model using CNN and Bi-directional LSTM in real-time environment. *Applied Sciences, 12*(8), 3989. https://doi.org/10.3390/app12083989

[13] Chowdhury, S. R., Khare, Y., & Mazumdar, S. (2023). Classification of diseases from CT images using LSTM-based CNN. In *Diagnostic Biomedical Signal and Image Processing Applications with Deep Learning Methods* (pp. 235–249). Elsevier. https://doi.org/10.1016/B978-0-323-91137-1.00014-0

[14] Hossain, M. M., Ali, M. S., Ahmed, M. M., Rakib, M. R. H., Kona, M. A., Afrin, S., Islam, M. K., Ahsan, M. M., Raj, S. M. R. H., & Rahman, M. H. (2023). Cardiovascular disease identification using a hybrid CNN-LSTM model with explainable AI. *Informatics in Medicine Unlocked, 42*, 101370. https://doi.org/10.1016/j.imu.2023.101370

[15] Moya, J. A. B. (2024). Addressing the gaps in early dementia detection: A path towards enhanced diagnostic models through machine learning. *arXiv preprint arXiv:2409.03147*. https://doi.org/10.48550/arXiv.2409.03147

[16] El-Sappagh, S., Alonso, J. M., Islam, S. M. R., Sultan, A. M., & Kwak, K. S. (2021). A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. *Scientific Reports, 11*(1), 2660. https://doi.org/10.1038/s41598-021-82213-5

[17] Saif, F. H., Al-Andoli, M. N., & Bejuri, W. M. Y. W. (2024). Explainable AI for Alzheimer detection: A review of current methods and applications. *Applied Sciences, 14*(22), 10121. https://doi.org/10.3390/app142210121

[18] Chen, K., Weng, Y., Hosseini, A. A., Dening, T., Zuo, G., & Zhang, Y. (2024). A comparative study of GNN and MLP based machine learning for the diagnosis of Alzheimer's Disease involving data synthesis. *Neural Networks, 169*, 442–452. https://doi.org/10.1016/j.neunet.2023.11.001

[19] Mirza, A. F., Mansoor, M., Usman, M., & Ling, Q. (2023). A comprehensive approach for PV wind forecasting by using a hyperparameter tuned GCVCNN-MRNN deep learning model. *Energy, 283*, 129189. https://doi.org/10.1016/j.energy.2023.129189

[20] Offodile, A. C., Gallagher, K., Angove, R., Tucker-Seeley, R. D., Balch, A., & Shankaran, V. (2022). Financial navigation in cancer care delivery: State of the evidence, opportunities for research, and future directions. *Journal of Clinical Oncology, 40*(21), 2291–2294. https://doi.org/10.1200/JCO.21.02221

[21] Munagandla, V. B., Dandyala, S. S. V., & Vadde, B. C. (2024). AI-powered cloud-based epidemic surveillance system: A framework for early detection. *Revista de Inteligencia Artificial en Medicina, 15*(1), 673–690. https://doi.org/10.1016/j.artmed.2023.101417

[22] Cohen, A. B., Stump, L., Krumholz, H. M., Cartiera, M., Jain, S., Sussman, L. S., Hsiao, A., Lindop, W., Ying, A. K., Kaul, R. L., & Others. (2022). Aligning mission to digital health strategy in academic medical centers. *NPJ Digital Medicine, 5*(1), 67. https://doi.org/10.1038/s41746-022-00611-4

[23] Addo, K. M., & Khan, H. (2024). Factors affecting healthy aging and its interconnected pathways. *Turkish Journal of Healthy Aging Medicine, 15*(1), 9–24.

[24] Newby, D., Orgeta, V., Marshall, C. R., Lourida, I., Albertyn, C. P., Tamburin, S., Raymont, V., Veldsman, M., Koychev, I., Bauermeister, S., & Others. (2023). *Artificial intelligence for dementia prevention. Alzheimer's & Dementia, 19*(12), 5952–5969. https://doi.org/10.1002/alz.13072

[25] Rabie El Kharoua. (2024). Alzheimer's disease dataset. *Kaggle*. https://doi.org/10.34740/KAGGLE/DSV/8668279