INJ INTERNATIONAL NEUROUROLOGY JOURNAL

Review and Comparison for Alzheimer's Disease Detection with Machine Learning Techniques

[1]Anil Chakravarthy, [2]B S Panda, [3]Sasmita Kumari Nayak
[1&3]Centurion University of Technology & Management, Odisha, India.
[2]Raghu Engineering College, Visakhapatnam, India.

Abstract

Alzheimer's disease is an advanced neurological condition marked by cognitive deterioration and memory loss. Early AD identification is crucial for quick treatment and better patient results. The diagnostic process can now be automated thanks to machine learning techniques, which can offer precise and effective tools for detecting early Alzheimer's disease. This study suggests an innovative method for detecting Alzheimer's disease that combines feature extraction, dimensionality reduction, and classification algorithms. The study makes use of a sizable dataset made up of neuroimaging scans, clinical evaluations, and demographic data from a wide range of people, including both people with confirmed diagnoses of Alzheimer's Disease and healthy controls. The study also investigates how various feature sets and mixtures of imaging modalities affect classification ability. In order to strengthen the robustness and generalization capacity of the Alzheimer's Disease detection model, an ensemble learning strategy is also being looked into.

An independent test set is used to validate the suggested approach, which is then thoroughly assessed using a cross-validation framework. The outcomes show how well the suggested method performs in precisely differentiating between those with Alzheimer's disease and healthy controls. Additionally, the model's performance is contrasted with other cutting-edge techniques for detecting Alzheimer's disease, emphasizing its competitive advantage in terms of precision and computational effectiveness. Algorithms for supervised classification are used in the suggested method. Classifiers such as Random Forests, Support Vector Machines and Navie base are used in a comparative analysis.

Key words: Alzheimer's Disease, Neurourology, Neuroimaging, Machine Learning, Feature Extraction, Dimensionality Reduction, SVM.

Introduction

The majority of dementia cases (50–60%) are caused by Alzheimer's disease (AD), a neurological disorder that normally develops gradually and gets worse with time [1]. The most common early symptom is difficulty remembering recent events [3]. A person with advanced Alzheimer's disease may experience difficulties with language, disorientation (including a propensity to get lost easily), mood fluctuations, a lack of drive, self-neglect, and behavioral issues. People commonly distance themselves from friends and relatives as their health deteriorates [4]. As body functions gradually decline, death is the end result. Three to nine years are typically the range of life expectancy after diagnosis, though the rate of advancement might vary [5].

The exact causation of Alzheimer's disease is unknown. Its development is influenced by a variety of genetic and environmental risk factors. The most potent genetic risk factor is an APOE allele [6]. Additional risk factors include a history of head trauma, severe depression, and high blood pressure. The illness process is largely influenced by amyloid plaques, neurofibrillary tangles, and a loss of neural connections in the brain. A probable diagnosis is obtained using the patient's medical history, cognitive testing, and blood tests to rule out any other potential reasons. Initial symptoms are typically mistaken for brain aging. Examining brain tissue is necessary for a specific diagnosis, but this can only be done after someone has passed away. In general, aging is known to be facilitated by a healthy diet, physical activity, and social engagement, which may also reduce the incidence of cognitive decline and Alzheimer's [7].

Although certain treatments may briefly lessen symptoms, none can stop or control the disease's course. Affected people become more and more dependent on others, which frequently puts a strain on carers. Social, psychological, physical, and economic factors can all play a role in the stress. Exercise routines may be beneficial for daily tasks and possibly improve outcomes [8]. Antipsychotics are frequently used to address

DOI: 10.5123/inj.2023.4.inj42
ISSN:2093-4777 | E-ISSN:2093-6931
Vol. 27 Iss. 4 (2023)

behavioral issues or psychosis brought on by dementia, however this is hardly advised because there is little benefit and a higher chance of dying young [9].

Around 50 million people will have Alzheimer's disease worldwide by the year 2020. The majority of cases affect adults over 65, although up to 10% of cases have an early onset and caused people in their 30s to mid-60s. About 6% of adults 65 and older are affected, and women are more often affected than males [10]. The illness bears Alois Alzheimer's name after the German pathologist and psychiatrist who originally characterized it in 1906 [11]. Alzheimer's disease places a heavy financial burden on society; it is thought to cost $1 trillion year globally. It is listed as the seventh most common cause of death in the US.

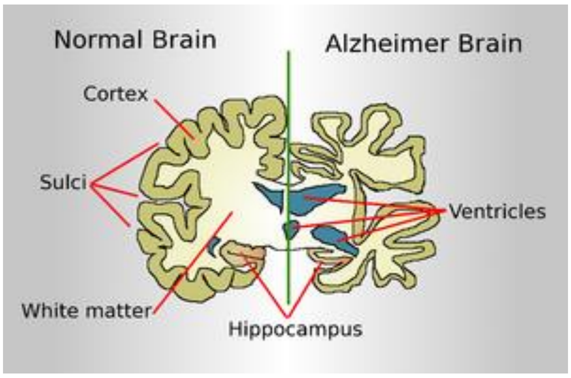

Figure 1: Alzheimer Disease

Literature Review

Numerous research has produced encouraging findings in recent years using artificial language processing to diagnose Alzheimer's disease (AD). We thoroughly examine these articles to determine the merits of this strategy, spot any problems, and present the key conclusions that can direct further investigation. In order to find relevant publications for this article, we used systematic literature reviews to search Ovid, PubMed, and Web of Science. Various methods [12] used to analyze brain scans in order to classify brain illnesses. Based on the symptoms of the evaluated studies, a number of important topics are covered in this literature.

Garrard et al. [13] demonstrated that bone-conducted speech could be constantly recorded and that perseverative patterns could be detected. We intend to correlate verbal repeat frequency with dementia stage, progression, and dementia type in next investigations. Scripts with incorporated repetitions were read by healthy participants. Principal Component Analysis was used to combine the features that were retrieved from the recorded signals and create a one-dimensional feature vector.

Sadeghian et al. [14] encourages the search for speech patterns that could indicate dementia. The cross-sectional design of this study, which aims to find a single pattern applicable to the entire range of subjects in each class, is one of its flaws. Each participant might act as his own control in a long-term study, reducing the significant within-group variation in speaking styles. The speech qualities that have been linked to dementia are not all included in the features used. A few more intriguing speech traits from the computational linguistics field could be added to our basket with little work. Given the small number of samples, the accuracy of 95% for detecting Alzheimer's disease appears encouraging.

Chien et al. [15] suggested a novel feature sequence model and used a data-driven method, namely the recurrent neural network, to perform classification in this study, in contrast to the majority of the previous literature, which retrieved statistics-based features and depended on a feature selection procedure. Additionally, the system is shown to be totally automated, indicating that it may be quickly installed everywhere. 150 speech samples were used in a series of studies to validate our findings, and the results showed an area under the receiver operating characteristic curve score as high as 0.825.

Konig et al. [16] studied that speech and language abnormalities are a symptom of several kinds of dementia and Mild Cognitive Impairment (MCI), which have been shown to be excellent predictors of the disease's existence and progression. Consequently, automatic speech analytics offered by a mobile application may be a helpful tool in providing further markers for assessment and identification of AD and dementia in their early stages.

ISSN:2093-4777 | E-ISSN:2093-6931
DOI: 10.5123/inj.2023.4.inj42
Vol. 27 Iss. 4 (2023)

Luz [17] suggested approach uses content-free characteristics collected from dialogical interaction in additive logistic regression, a machine learning boosting technique, to create a prediction model. The model training data included 25 conversations between interviewers and Alzheimer's patients and 18 conversations between interviewers and patients with different medical problems. Speech rate, turn-taking patterns, and other speech metrics were among the features examined. Our technique achieves overall accuracy of 86.5% despite depending just on content-free features.

Warnita et al. [18] offer a gated convolutional neural network (GCNN) method for automatically detecting Alzheimer's diseases using speech data. This GCNN can capture the temporal information in audio paralinguistic features and can be trained with a manageable quantity of data. It can be simply applied to any language because it does not make use of any linguistic traits. We used Pitt Corpus to assess our methodology. The proposed method outperformed the traditional sequential minimal optimization (SMO) by 6.8 points, achieving an accuracy of 73.6%.

Gosztolya et al. [19] unlike moderate cognitive impairment, which is clinically regarded as a prodromic stage of Alzheimer's disease (AD), which develops for years prior to clinical manifestation. Early diagnosis is essential for both types of neurodegenerative illnesses in order to enable prompt treatment and slow the disease's course. In this work, which builds on our earlier research, we attempted to distinguish between the three speaker groups by using automatically derived acoustic indicators from the subjects' spontaneous speech. the transcription of the spontaneous utterances was used to calculate several morphological, speech-based, and semantic linguistic aspects. But when we combined the two unique feature kinds, we achieved the best outcomes. One encouraging result of our investigation was that, compared to the scenario of employing all the linguistic features, the accuracy scores obtained only marginally fell when using only the semantic linguistic qualities.

Signs and Symptoms of Alzheimer's disease

During the course of Alzheimer's, the progression of cognitive and functional decline frequently occurs in three stages. Three stages are referred to as early or mild, middle or moderate, and late or severe. The disease is known to attack the hippocampus, which is linked to memory, and this causes the early signs of memory loss [20].

a. Early stage

A definitive diagnosis of Alzheimer's disease is made as a result of the disease's victims' increasing loss of learning and memory. In a small percentage of people, memory problems are less frequent than difficulties with language, executive functioning, perception (agnosia), or movement execution (apraxia). Alzheimer's disease does not necessarily have a negative impact on all memory abilities. The body's memory of how to carry out actions, such as using a fork to eat or a glass to drink from, is all influenced to a lesser amount than fresh facts or memories, according to episodic memory, semantic memory, and implicit memory [21].

b. Middle stage

Progressive deterioration finally makes it difficult for people to maintain their independence since they are unable to perform the majority of daily duties. Speech difficulties are indicated by frequent incorrect word replacements (paraphasias), which are caused by the inability to recall language. Furthermore, writing and reading skills degrade with age. Complex motor sequences become less coordinated over time as Alzheimer's disease develops, increasing the risk of falling. At this point, memory problems worsen, and it's possible that the person can no longer recognize their immediate family. One's long-term memory was normal before this happened.[22].

c. Final Stage

The final stage, often known as the late stage or the severe stage, is marked by complete dependence on caregivers. The usage of short phrases or even single words increases, eventually leading to speech loss. People typically understand and react to emotional signals even when their spoken language skills are compromised. Aggression is still a possibility, although excessive apathy and exhaustion are much more common symptoms. People with Alzheimer's disease progressively lose their muscle mass and mobility to the point that they are bedridden, unable to feed themselves, and incapable of performing even the most basic tasks independently. Typically, external problems like pneumonia or pressure ulcer infections rather than the illness itself are what lead to death. Just before dying, some persons encounter a paradoxical lucidity in which their mental clarity unpredictably reappears [23].

DOI: 10.5123/inj.2023.4.inj42
ISSN:2093-4777 | E-ISSN:2093-6931
Vol. 27 Iss. 4 (2023)

Figure 2: Stages of Alzheimer's Disease

Early signs and symptoms are:

- having trouble making decisions or solving problems.
- having trouble keeping up with conversations or finding words.
- having trouble performing routine tasks.
- forgetting things or recent events.
- losing or misplacing things.
- getting lost when walking or driving.
- being confused, even in familiar places.

Methodology

Alzheimer's disease is a complicated neurological condition that impairs daily functioning and causes progressive cognitive impairment. A prompt identification of AD is essential for effective treatment and better patient outcomes. Recent advances in computational approaches have made it possible to evaluate complicated neuroimaging data and clinical assessments, making machine learning techniques attractive tools for automating the AD identification process.

A. Neuroimaging Modalities in AD Detection:

The three main neuroimaging modalities used in the early detection of AD are structural MRI, functional MRI, and positron emission tomography (PET) studies. The extensive information provided by structural MRI regarding the brain's anatomy and the structural changes connected to AD. To extract useful features from structural MRI data, methods like area of interest (ROI) analysis and voxel-based morphometry (VBM) have been frequently used. Functional MRI records patterns of brain activity and connectivity that shed light on functional changes in AD. On the other hand, PET scans make it possible to see metabolic abnormalities related to AD pathology.

B. Feature Extraction and Representation:

An essential component of employing ML approaches to identify AD is feature extraction. To extract pertinent features from neuroimaging data, various techniques have been investigated. These consist of spatial domain features, statistical metrics, and texture analysis. Furthermore, cutting-edge approaches like feature extraction based on deep learning and graph-based representations have showed promise in capturing intricate patterns in brain images.

C. Dimensionality Reduction Techniques:

Because neuroimaging data is highly dimensional, dimensionality reduction techniques are used to simplify computation and enhance model performance. Principal Component Analysis (PCA), which efficiently captures the most informative features of the data, is still a well-liked technique for linear dimensionality reduction. Non-linear methods that can maintain local structure in high-dimensional data, such as t-distributed Stochastic Neighbor Embedding (t-SNE), have also gained popularity.

D. Classification Algorithms:

Algorithms for supervised classification are essential for separating AD patients from healthy controls. Among the most popular classifiers are Support Vector Machines (SVM), Random Forests, Neural Networks, and k-Nearest Neighbors (k-NN). By merging numerous base classifiers, ensemble approaches like AdaBoost and Random Forests have shown increased generalization abilities.

DOI: 10.5123/inj.2023.4.inj42
ISSN:2093-4777 | E-ISSN:2093-6931
Vol. 27 Iss. 4 (2023)

INJ INTERNATIONAL NEUROUROLOGY JOURNAL

E. Evaluation Metrics and Validation:
Area under the receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, accuracy, and other measures are used to evaluate the effectiveness of AD detection algorithms. To thoroughly assess model performance, cross-validation approaches like k-fold cross-validation and leave-one-out cross-validation are frequently used. The generalizability of the created models must also be evaluated by external validation on separate datasets.

F. Comparative Analysis and State-of-the-Art Approaches:
Several research have investigated how well various ML approaches work at detecting AD. To increase classification accuracy, some have focused on the integration of various neuroimaging modalities. Furthermore, investigating ensemble learning strategies and including clinical information alongside neuroimaging features has showed promise in boosting the stability of AD detection models.

Results & Discussions

A. Data collection:
The WHO created the Global Dementia Observatory (GDO), a data platform that compiles nation data on 35 important dementia indicators across the global action plan's seven strategic areas, to make it easier to monitor the global dementia action plan. The GDO Knowledge interaction Platform was introduced by WHO as an addition to the GDO with the aim of promoting reciprocal learning and multidirectional interaction across regions, nations, and individuals to support action globally. It is a repository of good practices examples in the field of dementia.

B. Data Preprocessing:
Data preparation is done to fill in blanks, eliminate unnecessary data, and fix mistakes. It is no longer required to include the sample code number in the data collection because it is irrelevant to the disorders themselves. 16 defining variables are absent from the dataset for whatever reason. The median makes up for the group's shortcomings. The dataset also uses random selection to ensure that the data are dispersed accurately.

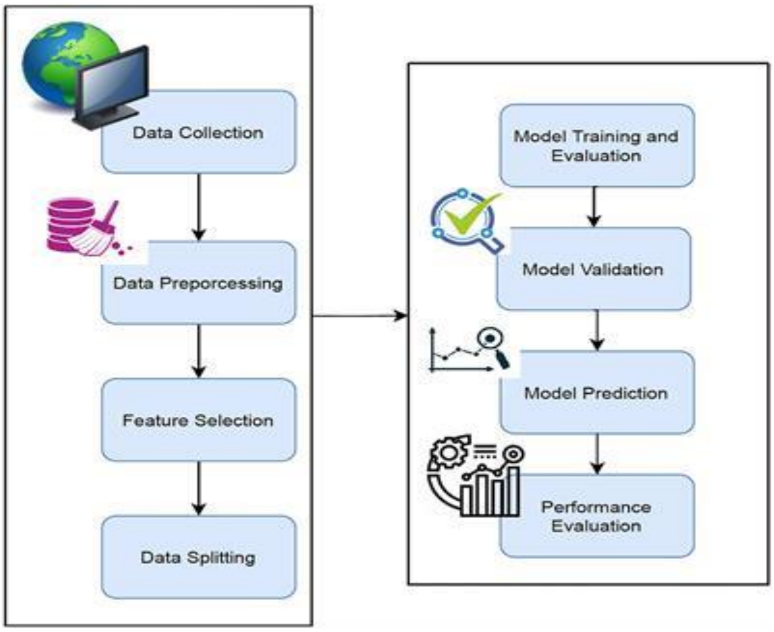

Figure 3: Alzheimer's Disease Prediction Model

C. Feature-Selection Methods:
The method of selecting the most trustworthy, non-redundant, and important features for a model is referred to as feature selection. It is essential to gradually minimize the size of datasets as their variety and quantity grow.

D. Data Splitting

DOI: 10.5123/inj.2023.4.inj42
ISSN:2093-4777 | E-ISSN:2093-6931
Vol. 27 Iss. 4 (2023)

Training-validation-test. The training, validation, and test sets are the three dataset divisions created by this data splitting approach. In this approach, the validation set is used to fine-tune the model to produce the best results while the training set is used to train the data.

E. Model Evaluation:
Model evaluation involves understanding a machine learning model's performance as well as its advantages and disadvantages using several evaluation measures, such as SVM, DT, NB, and Random Forest. Evaluating a model's performance is critical in the early stages of research. Model monitoring and assessment work together.

F. Model Prediction:
A popular statistical method for forecasting behavior is predictive modeling. Data-mining technology called predictive modeling solutions creates a model by studying past and present data and using it to forecast future results.

Author and Year	**Classification Techniques used**	**language and speech features**	**Performance Measure with results**
Ammer and Ayed 2018 [24]	SVM, NN, DT	MLU, POS, Repetition, word errors and morphemes	Precision = 79%
Chien et al 2018 [15]	RNN	Non-silence tokens and Speech length	AUC = 0.956
Gosztolya et al 2019 [19]	SVM	Acoustic attributes, semantic and morphological values	Acc = 86%
Hernandez-Dominguez et al 2018 [23]	SVM, Random Forest	Hapax legomena, information coverage and auxiliary verbs,	Acc = 87–94%
Konig et al 2018 [16]	SVM	Words' distribution in time and location of first word,	Acc = 86%
Luz 2018 [17]	NB	Speech rate, vocalization, number of utterances across discourse event	Acc = 68%

Table 1: Comparison Analysis of Alzheimer's Disease detection Using ML Techniques

Conclusion

The application of machine learning methods to the identification of Alzheimer's disease is an important development in the field of neurology. These methods provide strong instruments for early diagnosis, enabling prompt intervention and better patient outcomes. The capabilities of ML-based AD detection are being improved and expanded as part of ongoing research projects, with the ultimate goal of improving our knowledge of and ability to treat this deadly illness.

References:

1. *Ganguli M, Dodge HH, Shen C, Pandav RS, DeKosky ST (2005).* "Alzheimer disease and mortality: a 15-year epidemiological study". *Archives of Neurology. 62 (5): 779–84.*
2. *Burns A, Iliffe S (2009). "Alzheimer's disease". BMJ. 338:* doi:10.1136/bmj.b158.
3. *Knopman DS, Amieva H, Petersen RC, Chételat G, Holtzman DM, Hyman BT, et al. (2021).* "Alzheimer disease". *Nature Reviews Disease Primers. 7 (1): 33.* doi:10.1038/s41572-021-00269-y.PMC 8574196.
4. *Todd S, Barr S, Roberts M, Passmore AP (2013). "Survival in dementia and predictors of mortality: a review". International Journal of Geriatric Psychiatry. 28 (11): 1109–1124.* doi: 10.1002/ gps.3946.
5. *Long JM, Holtzman DM (2019).* "Alzheimer Disease: An Update on Pathobiology and Treatment Strategies". *Cell. 179 (2): 312–339.* doi:10.1016/j.cell.2019.09.001.
6. "Alzheimer's Disease Fact Sheet". *National Institute on Aging.* Archived *from the original on 23 March 2022.*
7. Dementia: assessment, management and support for people living with dementia and their carers *(Report).* National Institute for Health and Care Excellence *(NICE). 20 June 2018. NG97.*
8. "Low-dose antipsychotics in people with dementia". National Institute for Health and Care Excellence (NICE). Archived from the original on 5 December 2014. Retrieved 29 November 2014.

DOI: 10.5123/inj.2023.4.inj42
ISSN:2093-4777 | E-ISSN:2093-6931
Vol. 27 Iss. 4 (2023)

9. "Low-dose antipsychotics in people with dementia". National Institute for Health and Care Excellence (NICE). Archived from the original on 5 December 2014. Retrieved 29 November 2014.
10. Zhu D, Montagne A, Zhao Z (2021). "Alzheimer's pathogenic mechanisms and underlying sex difference". Cell Mol Life Sci. 78 (11): 4907–4920. doi:10.1007/s00018-021-03830-w. PMC 8720296. PMID 33844047.
11. Jump up to:a b c Berchtold NC, Cotman CW (1998). "Evolution in the conceptualization of dementia and Alzheimer's disease: Greco-Roman period to the 1960s". Neurobiology of Aging. 19 (3): 173–189. doi:10.1016/S0197-4580(98)00052-9. PMID 9661992. S2CID 24808582.
12. Panda, B. S., Satyabrata Patro, and Jyotirmaya Mishra. (2023) "Evaluation and Detection of Breast Cancer Using Data Mining Models." Machine Learning and Big Data Analytics: 2nd International Conference on Machine Learning and Big Data Analytics-ICMLBDA, IIT Patna, India, March 2022. Vol. 401. Springer Nature.
13. Garrard P, Nemes V, Nikolic D, et al. (2017) Motif discovery in speech: application to monitoring Alzheimer's disease. Curr Alzheimer's Res; 14 (9): 951–9.
14. Sadeghian R, Schaffer JD, Zahorian SA. (2017) Speech processing approach for diagnosing dementia in an early stage.
15. Chien YW, Hong SY, Cheah WT, et al. An assessment system for Alzheimer's disease based on speech using a novel feature sequence design and recurrent neural network. (2018) In: 2018 IEEE International Conference on Systems, Man, and Cybernetics; October 7–10, Miyazaki, Japan.
16. Konig A, Satt A, Sorin A, et al. Use of speech analyses within a mobile application for the assessment of cognitive impairment in elderly people. (2018) Curr Alzheimer's Res; 15 (2): 120–9.
17. Luz S, de la Fuente S, Albert P. A method for analysis of patient speech in dialogue for dementia detection. Audio Speech Process (2018); arxiv:1811.09919.
18. Warnita T, Inoue N, Shinoda K. Detecting Alzheimer's disease using gated convolutional neural network from audio data. Interspeech (2018) https://arxiv.org/abs/1803.11344 Accessed February 2,
19. Gosztolya G, Vincze V, Tóth L, et al. Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. (2019) Comput Speech Lang; 53: 1–17.
20. de Lira JO, Minett TSC, Bertolucci PHF, et al. Analysis of word number and content in discourse of patients with mild to moderate Alzheimer's disease. (2014) Dement Neuropsychol; 8 (3): 260–5.
21. Chien YW, Hong SY, Cheah WT, et al. An assessment system for Alzheimer's disease based on speech using a novel feature sequence design and recurrent neural network. (2018) In: 2018 IEEE International Conference on Systems, Man, and Cybernetics; October 7–10; Miyazaki, Japan.
22. Boyé M, Grabar N, Thi Tran M. Contrastive conversational analysis of language production by Alzheimer's and control people. (2014) Stud Health Technol Inform; 205.
23. Hernández-Domínguez L, Ratté S, Sierra-Martínez G, et al. (2018) Computer-based evaluation of Alzheimer's disease and mild cognitive impairment patients during a picture description task. Alzheimer's Dement; 10: 260.
24. Ammar R, Ayed Y. Speech processing for early Alzheimer disease diagnosis (2018) machine learning based approach. In: proceedings of IEEE/ACS International Conference on Computer Systems and Applications; October 28–November 1; Aqaba, Jordan.