International Journal of Scientific Research and Engineering Development-– Volume X Issue X, Year

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Dementia Predictor

Mrs. Supriya Suresh*, Ms. Meghana M**, Ms. Vaishnavi N Bhat***

*(Department of CSE, K S School of Engineering and Management, Bengaluru, India
Email: supriyasuresh@kssem.edu.in)

** (Department of CSE, K S School of Engineering and Management, Bengaluru, India
Email: meghanamahalinga410@gmail.com)

***(Department of CSE, K S School of Engineering and Management, Bengaluru, India
Email: vaishnbhat@gmail.com)

--************************----------------------------------

Abstract:

Dementia stands as a persistent and deteriorating ailment, posing a significant health challenge among older adults. Coping with the escalating cases of dementia presents a formidable task in the 21st century, particularly concerning patient care. The realm of Machine Learning (ML) applications, especially those geared towards "AI for social good," illustrates a compelling showcase of contemporary technological prowess. In this context, a digital biomarker for dementia emerges, aiming to anticipate the onset of this condition.

Data analysis constitutes a multifaceted process involving scrutiny, refinement, transformation, and modeling of data to unearth valuable insights, shape conclusions, and underpin decision-making. It encompasses a spectrum of techniques under various names, resonating across diverse domains encompassing business, science, and social sciences. Leveraging Machine Learning techniques facilitates predictive analytics, inherently intertwined with optimization, often cast as the minimization of loss functions over a training dataset.

The primary objective lies in furnishing a dementia prognosis based on an individual's data analysis, incorporating numerous attributes. Additionally, it proposes subsequent measures for afflicted individuals and extends guidance on preemptive measures. The dataset comprises patient information amassed subsequent to MRI scanning sessions. Analyzing an array of factors that influence outcomes through Support Vector Machine (SVM) and decision tree algorithms facilitates a deeper comprehension of their impact. This not only augments our understanding of diverse machine learning algorithms but also aids in selecting pivotal parameters to construct more robust models. Consequently, our model not only discerns the likelihood of dementia onset but also prescribes actionable steps for individuals without dementia to mitigate their risk.

Keywords — **Dementia, Machine Learning, Data Analytics, Support Vector Machine.**

--************************----------------------------------

I. INTRODUCTION

Recent advancements in integrated technologies within the medical sector have significantly enhanced the efficiency and precision of medical professionals. Among its various applications, disease prediction stands out as a primary focus. The landscape of healthcare procurement and management has witnessed substantial growth and evolution, particularly in improving health outcomes. In recent times, dementia has emerged as a prominent health concern among the elderly. Coping with dementia patients has become an increasingly formidable challenge in the modern era, given the persistent rise in dementia cases. Researchers have explored a multitude of techniques, including early-stage research, to detect dementia and its various stages, including Alzheimer's disease.

Furthermore, novel methodologies have surfaced for the early prognosis of dementia outcomes. The medical field's ongoing evolution has led to the establishment of vast, accessible databases for medical researchers. Despite this progress, accurately predicting disease progression remains a challenge for most patients. Hence, current endeavors focus on leveraging machine learning techniques to uncover and decipher patterns and relationships within extensive datasets. Through meticulous analysis, valuable insights are extracted to refine models supporting disease prognosis and enhance the accuracy of predicting patient health statuses.

Additionally, ongoing research is under scrutiny by various neuroscientists actively practicing in renowned hospitals such as Vani Villas and Trustwell Hospitals. These investigations are still ongoing, aiming to further advance our understanding and capabilities in disease prediction within the neurosciences domain.

II. TERMINOLOGIES

A. *Machine Learning*

The development of algorithms that allow computers to learn from data and improve their performance is called machine learning, a branch of artificial intelligence. Statistical, mathematical, and computer science techniques are used to examine and understand patterns in data. It is based on the idea of statistical learning. Many tasks such as natural language processing, predictive modeling, and image recognition can benefit from the application of machine learning techniques. They can learn from large datasets and adapt to new information, helping you make predictions and derive insights from complex data. Machine learning is used in many applications across industries such as manufacturing, healthcare, finance, and entertainment.

B. *Support Vector Machine*

Underpins Vector Machine, a administered machine learning calculation for classification and relapse errands. Amid classification, SVM finds the ideal hyperplane that best isolates information focuses from distinctive classes in a high-dimensional space. This hyperplane is chosen to maximize the edge, which is the remove between the hyperplane and the closest information focuses (bolster vectors) of each lesson. SVM can handle both directly divisible and non-linearly divisible information, utilizing diverse part capacities to outline the input information into higher dimensional spaces where the division is less demanding.

In relapse, SVM tries to discover a work that predicts the target variable by fitting a hyperplane as maybe closest to the information focuses. It minimizes the mistake between the anticipated values and the real values and penalizes the required deviation. SVM is known for its capacity to generalize well to inconspicuous information and its execution in high-dimensional spaces, making it a well-known choice for different machine learning tasks.

III RELATED WORK

A. *A Comparison of machine learning methods for survival analysis of high- dimensional clinical data for dementia prediction.[1]*

Clinical trials and cohort studies, particularly those in dementia research, often encompass intricate and multifaceted data. These datasets exhibit high dimensionality, indicating a vast array of variables or attributes relative to the number of observations available. Consequently, conventional statistical methodologies may prove impractical due to the extensive number of coefficients necessitating estimation. Moreover, the complexity of these datasets is compounded by potential instances of censorship, heterogeneity, and missing data. Addressing these challenges demands the development of techniques tailored to model and analyze such intricate data effectively. Furthermore, the stability of the models employed holds paramount importance, as it directly influences the reliability and confidence in the outcomes generated. The stability of an algorithm gauges its susceptibility to variations within the training dataset, wherein even minor alterations could yield significant fluctuations in the algorithm's performance. Hence, researchers are actively exploring methodologies to navigate these obstacles, aiming to furnish robust and steadfast analyses essential for informed decision- making within the realms of clinical research and healthcare provision.

B. *A machine learning approach for the differential diagnosis of Alzheimer and vascular dementia fed by MRI selected features.[2]*

Neurodegenerative diseases such as Alzheimer's disease (AD) witness the accumulation of amyloid-β plaques and tau-derived neurofibrillary tangles, mainly affecting the prefrontal and mesial-temporal regions of the brain. This pathological process causes significant shrinkage and atrophy of brain tissue and neurons, leading to memory loss and severe cognitive impairment. Accurate diagnosis is essential for effective management and therapy of AD. Advanced MR techniques, including diffusion tensor tomography (DTI) and resting-state functional magnetic resonance imaging (rs-fMRI), show promise in the diagnosis of dementia. By integrating machine learning (ML) techniques with MRI-based measures such as quantitative MRI, researchers have successfully distinguished AD patients from healthy elderly individuals. In addition, the synergistic application of ML and qMRI has shown potential to predict disease progression, including transition from mild cognitive impairment (MCI) to AD.

C. *Dementia prevention, intervention and care:2020 report of the Lancet Commission[3]*

Newest research shows that integrating three modifiable risk factors into the nine-factor life cycle model described by the 2017 Lancet Commission on Dementia Intervention, Prevention and Treatment can prevent or delay up to 40 percent of dementia cases. This highlights the importance of targeted interventions and public health initiatives, as prevention involves both individual actions and political implementation.

Implementation of specific measures such as maintaining systolic blood pressure at or below 130 mmHg in middle age, promoting the use of hearing aids, reducing . exposure to air pollution, preventing head injuries, limiting alcohol consumption, not smoking, ensuring universal access to primary and secondary education, reducing the risks of obesity and diabetes, and other potential risk factors such as sleep disorders are important steps throughout the lifespan.

In addition, address socio-economic inequalities is important and must provide support to people with dementia. In low and middle-income countries in particular, prioritizing preventive measures can lead to the most significant reductions in dementia cases.

D. *Toward a theory-based specification of non-pharmacological treatments in aging and dementia[4]*

Non-pharmacological treatments (NPTs) have the potential to be beneficial in treating clinical symptoms likewise also in primary and secondary prevention of dementia. NPTs offer numerous benefits, including widespread acceptance, few adverse effects, and significant and simultaneous compatibility with pharmaceutical treatments and other NPTs without significant interference problems. The various clinical stages of the illness, such as dementia, moderate cognitive impairment (MCI), and even cognitively normal people who are at danger of dementia, can all be treated with NPTs. So, as age-related neurodegenerative disorders progress, NPTs may have a significant effect on wellbeing, cognition, and quality of life. NPT offers several different types of interventions, such as dietary therapies, physical therapy, cognitive training, art therapy, and memory therapy.

An influential Previous Meta-Analysis defined NPTs as "any theoretically based, nonchemical, focused, and replicable intervention, conducted with the patient or the carer, which potentially provided some relevant benefit". The low quality of the majority of the data supporting multiple NPTs has been consistently highlighted by this and numerous previous comprehensive examinations of NPTs in dementia and ageing. To raise the calibre of the evidence supporting NPTs, more focused, reliable, and well-reported research likewise stricter methodological guidelines are required. This will assist in validating the extent and upper bounds of the advantages connected to various NPT kinds

E. *Statistical methods for dementia risk prediction and recommendations for future work: A systematic review[5]*

Over the final decade, numerous calculations have been created to anticipate the chance of dementia and recognize chance bunches. Three of his meta-analyses and efficient surveys summarizing dementia chance forecast models have been distributed in later a long time. After getting the execution examination, the affectability, specificity and region beneath the bend of distributed dementia chance models were inspected. The most center was on evaluating the factors included within the show and its prescient execution. It was concluded that none of the distributed models can be suggested to anticipate the risk of dementia due to methodological imperfections within the inquire about strategies and the models utilized to construct the models. The issue of the appraisal interims of hazard bunches, the failure to recognize between dementia subtypes, the need of outside and inner approval of models and, over all, the vulnerability of the investigation strategies utilized, were too pointed out. Methodological mistakes recognized within the demonstrate, such as need of outside approval and deficiently prescient control (region beneath the bend >= 0.74) in numerous populaces, counting diabetic, middle-aged and elderly patients, were also considered.

IV DATASET

The dataset that has been used in this paper was provided by the Open Access Series of Imaging Studies (OASIS) project which aims at making MRI data sets of the brain freely available for future discoveries. The chosen type of dataset is Longitudinal MRI data in nondemented and demented older adults. Longitudinal MRI data has the following attributes:

- This dataset consists of a longitudinal collection of 150 subjects aged 60 to 96.
- Each subject was scanned on two or more visits.
- There are both men and women and everyone is right-handed.
- 72 subjects were characterized as ‘nondemented’.
- 64 subjects were characterised as ‘demented’.
- 14 subjects were grouped as ‘Nondemented’ at the time of their visit, who later were characterised as ‘Demented’ on their following visits. These fall under the ‘Converted’ category.

V ARCHITECTURE

The study focused primarily on putting into practice the proposed architecture and developing a functional model suitable for real- time application. Following thorough research and a review of existing literature, a comprehensive architectural design for the proposed model was formulated. Illustrated in Figure 1, this architecture consists of sequence of steps, each requiring diligent implementation and comprehension of its sub-stages.

The initial step involves data preparation, which entails gathering MRI data and readying it for subsequent processing. Data collection holds paramount importance as it sets the groundwork for the entire process. Subsequently, the second step encompasses data pre-processing tasks, where set of attributes are taken into consideration and if those attributes are out of range they have been ignored in the following step. This stage addresses various issues such as handling missing data, grouping age-related features, removing outliers, and standardizing features. Visual representations such as charts and histograms aid in understanding the raw data, playing a crucial role in shaping the accuracy of subsequent stages.

The third stage focuses on data segmentation, specifically dividing the dataset into training and testing subsets. These segregated datasets are then utilized in the subsequent stage of model building, which includes sub-stages like model training, evaluation, and cross-validation. This phase is dedicated to implementing machine learning

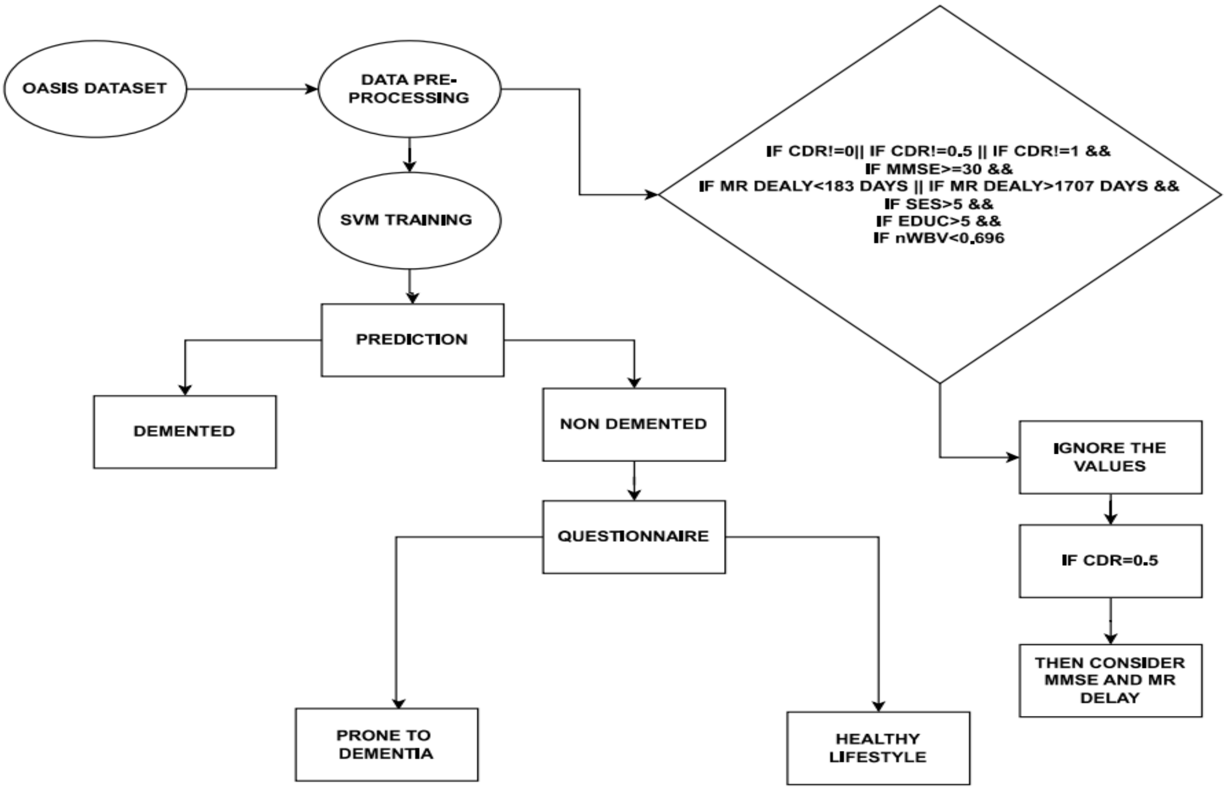

Fig 1 Architecture of the model

algorithms, training various classifiers, and evaluating model performance through accuracy assessments and cross- validation techniques. Model evaluation entails the application of multiple machine learning algorithms for data learning and classification, resulting in model generation.

Finally, the last step involves result generation, comprising model prediction and evaluation based on previously developed models. The generated results undergo evaluation using performance metrics, visually depicted to offer insights into model performance.

VI ALGORITHM

Step 1. Start
Step 2. Import necessary libraries: Flask, render_template, request, redirect, url_for, pandas, pickle.
Step 3. Initialize the Flask application.
Step 4. Load the trained model from the pickle file.
Step 5. Define a function to preprocess input data.
Step 6. Define routes for different pages:
 a. Route for the home page
 b. Route for the results page
 c. Route for the question page
Step 7. Define the function for the home page
 a. If request method is POST:
 b. Get user input from the form.
 c. Create a data frame with the user input.
 d. Preprocess the user input.
 e. Make prediction using the trained model.
 f. If prediction is 0, redirect to 'results' page.
 g. Otherwise, render the 'result.html' template with the prediction.
 h. If request method is GET, render the 'index.html' template.
Step 8. Define the function for the 'results' page
 a. Redirect to 'question' page.
Step 9. Define the function for the 'question' page
 a. Render the 'question.html' template.
Step 10. Run the Flask application.
Step 11. Stop

VII CONCLUSIONS

An exploration on the methods used for the prediction and analysis of dementia can be done using various factors and features. All these prediction methods mostly include algorithms of machine learning. Using the Support Vector

Machine algorithm a probable accuracy of 91.2% can be obtained.

ACKNOWLEDGMENT

The authors of the paper completely acknowledge the information that is presented in this paper and as well as all the contents present. For the details of the authors please refer the top of the page.

REFERENCES

[1] [1] Annette Spooner, Emily Chen , Arcot Sowmya , Perminder Sachdev, Nicole A. Kochan , JulianTrollor and Henry Brodaty. A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction,(2020).

[2] [2] Gloria Castellazzi , Maria Giovanna Cuzzoni , Matteo Cotta Ramusino , Daniele Martinelli , Federica Denaro , Antonio Ricciardi , Paolo Vitali , Nicoletta Anzalone, Sara Bernini, Fulvia Palesi , Elena Sinforiani , Alfredo Costa , Giuseppe Micieli, Egidio D'Angelo, Giovanni Magenes and Claudia A. M. Gandini Wheeler-Kingshott. I. A machine learning approach for the differential diagnosis of Alzheimer and vascular dementia fed by MRI selected features.(2020).

[3] [3]Gill Livingston, Jonathan Huntley, Andrew Sommerlad, David Ames, Clive Ballard, Sube Banerjee, Carol Brayne, Alistair Burns, Jiska Cohen-Mansfield, Claudia Cooper, Sergi G Costafreda, Amit Dias, Nick Fox, Laura N Gitlin, Robert Howard, Helen C Kales, Mika Kivimäki, Eric B Larson, Adesola Ogunniyi, Vasiliki Orgeta, Karen Ritchie, Kenneth Rockwood, Elizabeth L Sampson, Quincy Samus, Lon S Schneider, Geir Selbæk, Linda Teri, Naaheed Mukadam. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission.(2020).

[4] [4] Sietske A.M. Sikkes, Yi Tang, Roos J. Jutten, Linda M.P. Wesselman, Lyn S. Turkstra, Henry Brodaty, Linda Clare, Erin Cassidy- Eagle, Kay L. Cox, Gaël Chételat, Sophie Dautricourt, Klodian Dhana, Hiroko Dodge, Rose-Marie Dröes, Benjamin M. Hampstead, Thomas Holland, Amit Lampit, Kate Laver, Antoine Lutz, Nicola T. Lautenschlager, Susan, M. McCurry, Franka J. M. Meiland, Martha Clare Morris, Kimberly D. Mueller, Ruth Peters, Gemma Ridel, Aimee Spector, Jenny T. van der Steen, Jeanette Tamplin, Zara Thompson, Alex Bahar-Fuchs. Toward a theory-based specification of non- pharmacological treatments in aging and dementia: Focused reviews and methodological recommendations.(2020).

[5] [5] Jantje Goerdtena, Iva Cukic , Samuel O. Dansoa , Isabelle Carriereb , Graciela Muniz- Terrera. Statistical methods for dementia risk prediction and recommendations for future work: A systematic review.(2019).

[6] Tomasz M. Rutkowski, Masato S. Abe, and Mihoko Otake-Matsuura. Neurotechnology and AI Approach for Early Dementia Onset Biomarker from EEG in Emotional Stimulus Evaluation Task.(2021)

