DETECTION OF NEUROCOGNITIVE IMPAIRMENT IN THE ELDERLY USING DEEP LEARNING

PHASE I REPORT

Submitted by

OVIYA K	**210701186**
NAVYA BALASUNDARAM	**210701177**

in partial fulfilment for the award of the degree of

BACHELOR OF ENGINEERING
in
COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

ANNA UNIVERSITY, CHENNAI 600 025

NOVEMBER 2024

RAJALAKSHMI ENGINEERING COLLEGE

BONAFIDE CERTIFICATE

Certified that this report titled "**DETECTION OF NEUROCOGNITIVE IMPAIRMENT IN THE ELDERLY USING DEEP LEARNING**" is the bonafide work of "**OVIYA K (210701186), NAVYA BALASUNDARAM (210701177)"** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE	**SIGNATURE**
Dr.P.Kumar., M.E, PhD	**Dr. Ananthajothi K, PhD**
Head of Department	**Professor, Supervisor,**
Department of Computer Science and Engineering,	Department of Computer Science and Engineering,
Rajalakshmi Engineering College,	Rajalakshmi Engineering College,
Chennai - 602105	Chennai – 602105

Submitted to Project Viva-Voce Examination held on___________.

Internal Examiner **External Examiner**

ACKNOWLEDGEMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavour to put forth this report. Our sincere thanks to our Chairman **Mr. S. MEGANATHAN, B.E, F.I.E.**, our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN**, **Ph.D.,** for providing us with the requisite infrastructure and sincere endeavouring in educating us in their premier institution. Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. P. KUMAR, Ph.D.**, Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guide, **Dr. ANANTHAJOTHI K, PhD.,** Department of Computer Science and Engineering. Rajalakshmi Engineering College for her valuable guidance throughout the course of the project. We are very glad to thank our Project Coordinator, **Dr. KUMARAGURUBARAN T, Ph.D.** Department of Computer Science and Engineering for her useful tips during our review to build our project.

OVIYA K (210701186)

NAVYA BALASUNDARAM (210701177)

ABSTRACT

As the global population ages, neurocognitive disorders like the Alzheimer's disease which is a form of dementia, and several other forms of dementia are becoming more commonplace. These conditions significantly impair the quality of life for patients and present a growing challenge for healthcare systems worldwide. Early detection is critical to improving patient outcomes, yet traditional diagnostic methods often struggle to catch on to the very early and subtle signs of decline in cognitive function. In this paper, a deep learning algorithm, TabNet, is used on clinical data to detect neurocognitive impairment. Before applying the algorithm, Exploratory Data Analysis is also performed to get a better understanding of the dataset. This research is mainly focused on people with family history. The TabNet model, designed especially for tabular data easily and effectively learns the importance of each feature and their inter-dependencies. In future work, it is aimed to make this research more holistic by using Vision Transformer models to analyze MRI images. This study currently focuses on clinical data, which includes demographic, lifestyle, and cognitive assessments, and demonstrates how TabNet can be used to identify signs of cognitive impairment. Working with critical features, this study also focuses on data pre-processing and feature selection. This project's main aim is to utilize AI in healthcare to help with diagnoses that are otherwise somewhat difficult to make.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATONS	**ACRONYMS**
EDA	EXPLORATORY DATA ANALYSIS
MMSE	MINI MENTAL STATE EXAMINATION
MRI	MAGNETIC RESONANCE IMAGING
EHR	ELECTRONIC HEALTH RECORDS
SCA	SMART CONVERSATIONAL AGENT
MCI	MILD COGNITIVE IMPAIRMENT

CHAPTER 1
INTRODUCTION

1.1 GENERAL

Neurocognitive disorders, including Alzheimer's, Parkinson's, and various forms of dementia, are increasingly common as global life expectancy rises. These disorders often begin with subtle cognitive symptoms that can be difficult to detect using traditional diagnostic tools. Unfortunately, by the time these symptoms are noticeable enough for standard evaluations to identify them, substantial brain damage may already have occurred. As a result, patients are often diagnosed later in life, which limits the effectiveness of available treatments and lowers the chances of successful intervention. Early detection is crucial for improving patient outcomes, as it can facilitate timely interventions that reduce symptom severity, slow disease progression, and enhance the quality of life. In this context, artificial intelligence (AI), particularly deep learning, offers significant potential. AI can analyze vast amounts of clinical data to uncover hidden patterns that may not be immediately obvious to human clinicians. This ability is especially valuable in detecting neurocognitive disorders, where early diagnosis is essential for improving patient outcomes.

This study focuses on using the TabNet deep learning algorithm to predict neurocognitive disorders in elderly patients, utilizing clinical data that includes demographic information, lifestyle factors, medical history, and cognitive assessments. TabNet is chosen for its ability to handle tabular datasets effectively and its capacity to dynamically learn which features are most important for prediction. The goal is to train the TabNet model to accurately recognize the early stages of cognitive impairment, allowing for more timely and precise detection compared to conventional diagnostic methods.

As TabNet learns from the features in the dataset, it can identify dependencies between them, refining the accuracy of the model. TabNet compares each feature with every other feature precisely and learns about their interdependence.

In future research, MRI images will be analyzed using the Vision Transformer model, which will provide additional structural insights into brain health, further enhancing the predictive power of the system. By integrating imaging data with the clinical data used in this study, the diagnostic tool will offer a more comprehensive view of a patient's condition. To ensure the best results, the research includes an exploratory data analysis (EDA) phase to better understand the features within the dataset, which will inform the TabNet model and improve its accuracy in diagnosing neurocognitive disorders.

1.2 OBJECTIVE

The primary objective of this project is to develop a deep learning model that accurately predicts signs of neurocognitive impairment in the elderly using routine medical check-up data. The system aims to provide healthcare providers with a tool that facilitates diagnosis, allowing for timely interventions and better management of neurocognitive disorders. This detection will prevent late diagnosis which will be a ineffective if the disease has already progressed past a threshold. For this project, relevant patient information will be utilized effectively to provide accurate prediction results.

This project also aims to bridge the gap between traditional diagnostic methods and modern technological advancements by leveraging AI to enhance predictive accuracy. By integrating routine clinical data with advanced algorithms, the system not only improves the reliability of early detection but also makes the process more accessible and cost-effective for healthcare providers. Unlike resource-intensive approaches such as imaging or invasive procedures, this model relies on easily obtainable patient data, enabling widespread implementation in both urban hospitals and rural clinics. This democratization of diagnostic tools ensures that more patients receive timely evaluations, improving outcomes and reducing the overall burden on healthcare systems. By focusing on clinical data that is readily available, the project aims to minimize diagnostic delays while reducing dependency on costly imaging techniques.

This system aspires to serve as a scalable solution, adaptable to various healthcare settings, improving early detection rates and enabling timely interventions for better patient outcomes.

1.3 EXISTING SYSTEM

Existing systems for detecting neurocognitive disorders primarily rely on traditional diagnostic approaches such as cognitive assessments, including the Mini-Mental State Examination (MMSE), and neuroimaging techniques like Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans. These methods, while accurate in diagnosing advanced stages of disorders, often fail to detect subtle early signs of cognitive decline. As a result, many patients are diagnosed at later stages when therapeutic interventions are less effective. Additionally, these methods are resource-intensive, requiring specialized equipment and expertise, which limits their accessibility, especially in rural or resource-constrained healthcare settings. The reliance on time-consuming manual evaluations further adds to the delays in diagnosis, exacerbating the issue of late-stage detection.

Another major limitation of existing systems is their inability to integrate diverse patient data for holistic analysis. Most diagnostic tools focus on individual factors, such as isolated cognitive scores or single medical tests, rather than leveraging the interplay of multiple health parameters. This fragmented approach fails to capture the complex relationships between factors like family medical history, lifestyle choices, and cognitive assessments, which are crucial for identifying early neurocognitive impairments. Consequently, potential early indicators are often overlooked, leaving healthcare providers with an incomplete understanding of a patient’s condition.

In addition to these issues, the existing systems often lack predictive capabilities. While they excel at confirming diagnoses once symptoms are apparent, they provide little insight into the likelihood of future cognitive decline based on current health data. This limitation prevents healthcare providers from taking proactive measures to slow disease progression or mitigate risk factors. The absence of predictive analytics also makes it difficult to personalize treatment plans or prioritize patients who are most likely to benefit from early intervention.

1.4 PROPOSED SYSTEM

The proposed system addresses the limitations of traditional diagnostic approaches by leveraging a deep learning model, specifically TabNet, to predict neurocognitive impairments based on routine medical check-up data. Unlike conventional systems, this model integrates demographic, medical history, lifestyle, and cognitive assessment data into a single analytical framework. TabNet's ability to dynamically select the most relevant features ensures that key indicators, such as family history of Alzheimer's or MMSE scores, are prioritized during analysis.

One of the major strengths of the proposed system is its focus on routine clinical data, which is easily obtainable during standard check-ups. By utilizing features such as age, blood pressure, cholesterol levels, and behavioral symptoms, the system eliminates the need for specialized equipment or advanced diagnostic tests in the initial stages. This makes the solution highly adaptable to both urban hospitals and resource-constrained rural clinics, democratizing early detection capabilities. Additionally, the system is designed to work seamlessly with existing electronic health records (EHRs), ensuring smooth integration into hospital workflows.

The proposed system incorporates a streamlined data preprocessing pipeline to handle missing values, normalize data, and encode categorical variables. This ensures that the input data is clean and consistent, which is critical for reliable predictions. Exploratory Data Analysis (EDA) forms an integral part of the system, helping identify patterns, correlations, and outliers that may influence the model's performance. By combining EDA with TabNet's advanced feature selection, the system is able to extract meaningful insights from complex datasets, uncovering hidden relationships between variables that are often overlooked in traditional approaches.

CHAPTER 2
LITERATURE SURVEY

This paper systematically reviews the use of SCAs [2] for neuropsychiatric disorder detection. It identified 17 studies based on disorders like depression, dementia, and other impairments in cognition, by using AI technologies in feature extraction and analysis through deep learning methods. Most systems that are related to verbal and audiovisual data for prediction validation lack comprehensiveness, especially as regards reliability and usability. Though the field holds great promise and is increasingly integrating AI, current SCAs are not yet validated for clinical use.

Predictive models were created in a different study [10] to evaluate the change from mild to serious neurocognitive dysfunction. Over a two-year period, 132 subjects' clinical, demographic, and neuropsychological data were incorporated into the model. It identified important risk variables such as high body mass index and alcohol intake with an accuracy of 83.7%. The danger was decreased by characteristics like being female and having better praxis ability. The generalizability of the model was constrained by the small dataset.

By combining a Support Vector Machine (SVM) with a Feature Extraction Battery (FEB) [8], the proposed FEB-SVM algorithm sought to enhance dementia prediction. The model outperformed 12 state-of-the-art techniques, reaching accuracy of 98.28% on training set and accuracy of 93.92% on testing set. FEB's created features were successful in increasing prediction accuracy, but they were ineffective in pinpointing the precise causes of dementia. More feature development was stressed throughout the study in order to gain deeper insights.

Population and care data were merged in a Swedish longitudinal study (SNAC) [13] to record the aging process and care requirements in various locations. Many aspects of senior health, care, and social services were examined in this study.

The multidisciplinary methodology made it possible to analyze aging holistically. Unfortunately, the extensive breadth necessitated complicated and resource-intensive data collection and study design. The article emphasized the difficulties of overseeing a dataset this size.

In a study of Saudi patients, an AI-based method [12] for diagnosing neurocognitive problems that employed logistic regression and SVM models achieved up to 95.5% accuracy. Important variables like chronic illnesses and educational attainment were brought to light by the study. Even though the SVM model worked well, more research was needed to determine whether it could be applied to other populations. The study found that in order to improve generalizability, it is crucial to use a variety of datasets.

Using 2D MRI data, a Bi-Vision Transformer (BiViT) model [4] provided a novel transformer-based method for classifying cognitive diseases, including Alzheimer's disease. The model used cutting-edge feature learning modules to achieve 96.38% accuracy. It captured intricate patterns in medical photos better than previous deep learning techniques. Its application in real-world scenarios is constrained in the absence of more data due to its poor performance in smaller or unbalanced datasets.

An analysis of AI-augmented neurocognitive screening tests [5] examined the ways in which deep learning and speech recognition are examples of AI and machine learning techniques that have enhanced cognitive evaluations. The study covered how AI might improve exam accuracy and lessen biases. Notwithstanding these developments, issues with technological accessibility, data privacy, and dependability nonetheless exist. It was noted that older-friendly designs and more validation were required.

By merging weak learner models, the StackEnsembleMind [15] model improved mental health assessments through a stack-based ensemble machine learning technique. The model identified mental states including stress and worry with 98% accuracy.

Reliability was increased by addressing class imbalances with SMOTE. Unfortunately, the model's intricacy limited its practical application by making it challenging to understand and apply in healthcare settings.

Another paper explores the application of machine learning, particularly deep learning [11] , for the classification of three neurodegenerative diseases: cerebral ataxia, Alzheimer's disease, and Parkinson's disease. The authors uses a two-step methodology, employing pre-trained neural network architectures in the form of VGG16 and ResNet101. In fine-tuning these models for each disease, the researchers enhance the networks' ability to differentiate among them. Optimizations techniques like SGD, Rmsprop, and Adam are used to get the best result; evaluation metrics such as accuracy, loss are applied for the assessment of the model. The model is deployed through the Flask web framework, making it accessible and user-friendly

A paper discusses the use of machine learning algorithms in disease diagnosis, focusing on the application of these algorithms in Electronic Health Records (EHRs) [3]. Because of the ever-expanding patient data, machine learning is promising for improving the efficiency and accuracy of diagnosis. The study uses anonymized EHRs to include the medical history, laboratory results, and clinical notes to build models for predicting diseases. Various ML algorithms were implemented and evaluated against their interpretability, computational efficiency, and diagnostic accuracy. The results show that machine learning-based methods are superior compared to conventional diagnostic methods, providing improved accuracy and efficiency in disease diagnosis and prediction.

In order to predict mild cognitive impairment (MCI) [7] based on emotional arousal and valence tasks, a study that classified MCI from behavioral responses using machine learning. The model's accuracy in differentiating between MCI and normal cognition was around 90%. Comparing this non-invasive technology to traditional ones revealed practical advantages.

To capture more sophisticated signs of cognitive impairment, the fundamental components of the activities needed to be further refined.

Analysis of vocal characteristics [14] such as speech rate and voice activity in natural conversations was the main goal of a long duration study about the use of voice data for Alzheimer's disorder detection. The technique used a Bayesian classifier and obtained 68% accuracy. Non-invasive speech data was a promising early diagnosis technique because of how easy it was to use. The quantity of the dataset, however, hampered its accuracy, indicating the need for further data to enhance performance.

A machine learning-based dual-task gait assessment model [6] was developed to detect Alzheimer's disorder and mild to medium cognitive decline. When the DSM-5 criteria were compared to those of other diagnostic systems, the model discovered that DSM-5 detected more cases than DSM-IV. Still, diagnosis accuracy for less severe illnesses like MCI was only moderately high. Although more work needed to be done, the study indicated that the DSM-5 criteria could help with early cognitive decline identification.

MRI data was utilized to build a 2D Convolutional Neural Network (CNN) [1] to identify moderate cognitive impairment (MCI) at early stage, which produced a user-friendly diagnostic tool. The model showed promise for application in resource-constrained clinical settings and performed well even with little datasets. But because of system constraints, crucial preprocessing stages like skull stripping were omitted, which might have affected performance.The study made recommendations for future developments in data preparation.

Alzheimer's disease was diagnosed using a multimodal technique that included gene expression data and MRI images, using deep learning models like CNN and SpinalNet [9] A complete method of diagnosis was made possible by the combination of genetic and imaging data. However, the method's processing cost restricted its full capacity areas with insufficient resources.

The study underlined the need for more effective implementations while also highlighting the possibilities of multimodal techniques

A paper proposed a system that aims to detect Parkinson's Disease (PD) [16] at an early stage using Magnetic Resonance Imaging (MRI) and artificial intelligence. MRI is very essential in identifying the characteristics of Parkinson's features, and this study makes use of deep learning, more importantly, a Deep Ensemble Network, to enhance its detection process. The combined system uses a CNN model and LSTM optimized using the ADAM algorithm. The accuracies, precisions, recalls, and F1 values obtained are used to depict the performance of the proposed model, which attains an astonishing 98% accuracy. Comparison with existing approaches proves that the proposed method ensures effective early detection of PD.

CHAPTER 3
SYSTEM DESIGN

3.1 GENERAL

3.1.1 ARCHITECTURE DIAGRAM FOR EDA

The figure 3.1.1 shows the process of EDA. Data preparation starts with the system architecture. The clinical dataset is cleaned and standardized. Missing values for continuous variables were filled by Mean. Continuous features like BMI are normalized, which reduces the effect of outliers. All the data that is unnecessary was filtered out through feature engineering after preprocessing, leaving only the features useful in the prediction of neurocognitive disorders.

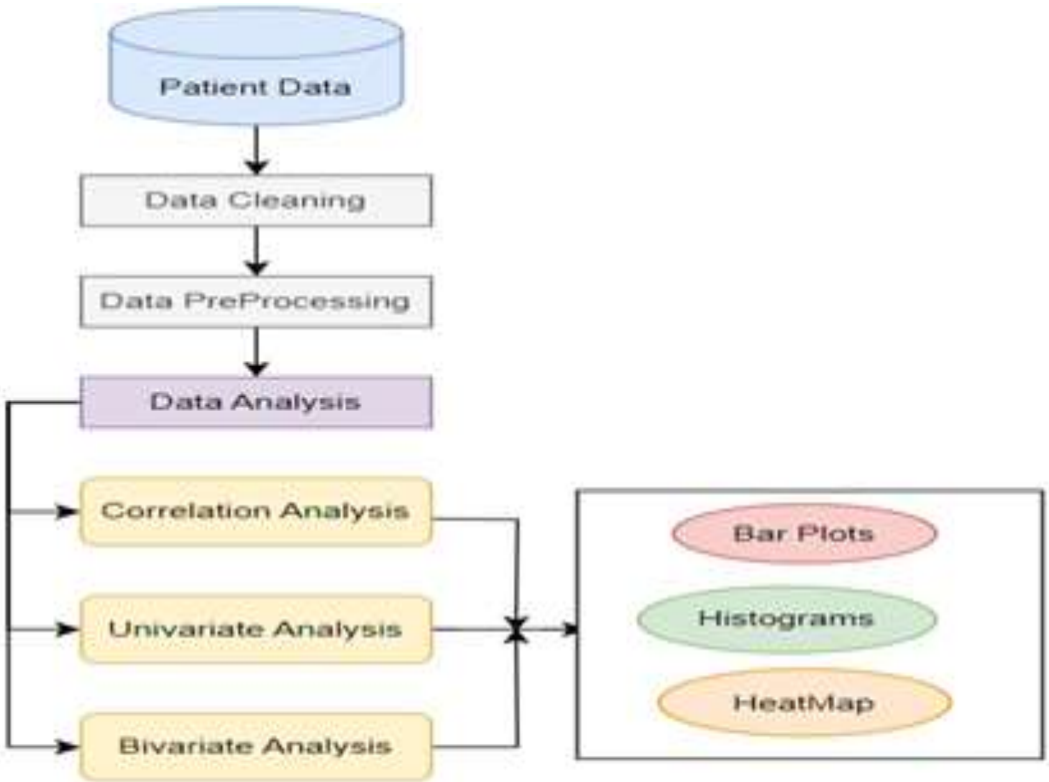

Fig 3.1.1.: System Architecture for EDA

3.1.2 ARCHITECTURE DIAGRAM FOR TABNET

TabNet also has the ability to handle clinical data by transforming the raw data features with the feature transformer layers. These layers are applied to create more sophisticated representations, allowing the model to understand complex interactions among variables. For example, Age, SystolicBP, and CardiovascularDisease are feature interactions that influence the score of cognitive health all at once and are computed to highlight their interaction effect on neurocognitive disorders. The raw inputs being transformed to high-dimensional representations here enable deeper discovery of subtle dependencies relevant to making precise predictions.

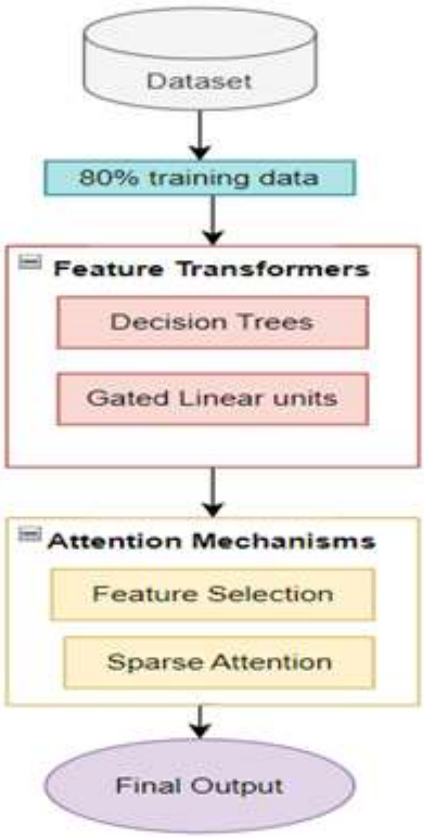

Fig 3.1.2 Working of TabNet Algorithm

3.1.3 USE CASE DIAGRAM

The Use Case Diagram illustrates the interactions between the key actors—Patient and Clinician—and the Neurocognitive Disorder Detection System. The Patient provides clinical data, such as demographics and cognitive assessments, to the system. This data is processed through various modules, including data preprocessing, exploratory data analysis (EDA), and model training. The Clinician reviews the system's predictions and insights to make informed decisions for diagnosing and managing neurocognitive disorders. The diagram clearly shows how these interactions help facilitate early diagnosis and timely interventions.

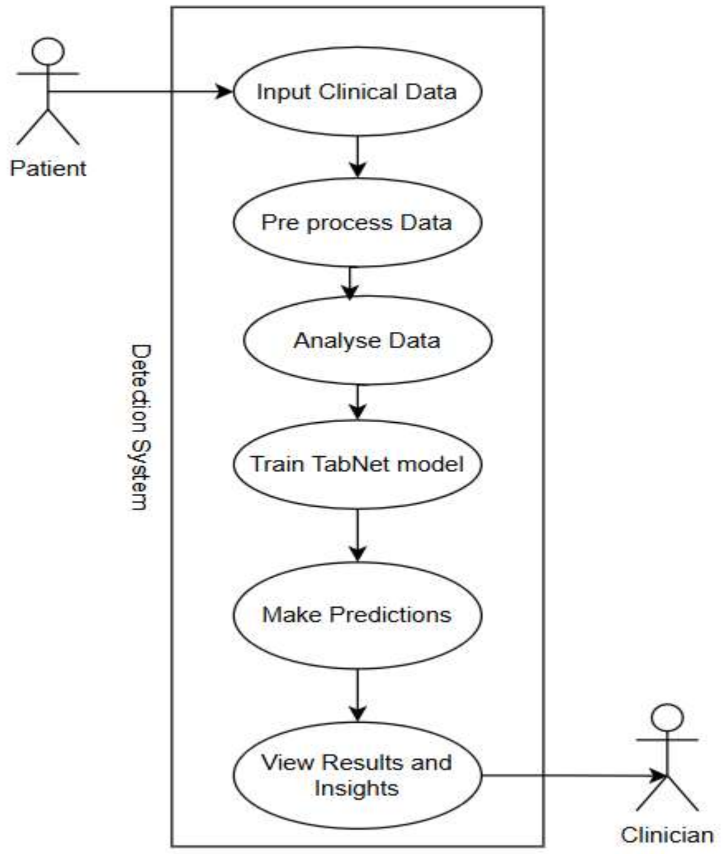

Fig 3.1.3 Use case Diagram

3.1.4 SEQUENCE DIAGRAM

The Sequence Diagram depicts the step-by-step flow of actions within the system. It begins with the Patient providing clinical data, which is then passed to the Input Data Module. After preprocessing, the data is analyzed by the EDA module, and insights are forwarded to the Model Training Module. Once the model is trained, it is used to generate predictions, which are provided to the Clinician. Additionally, feedback may be given to the Patient. This diagram showcases the dynamic process from data collection to prediction delivery, emphasizing the interactions and sequence of events required to reach actionable insights for the Clinician.

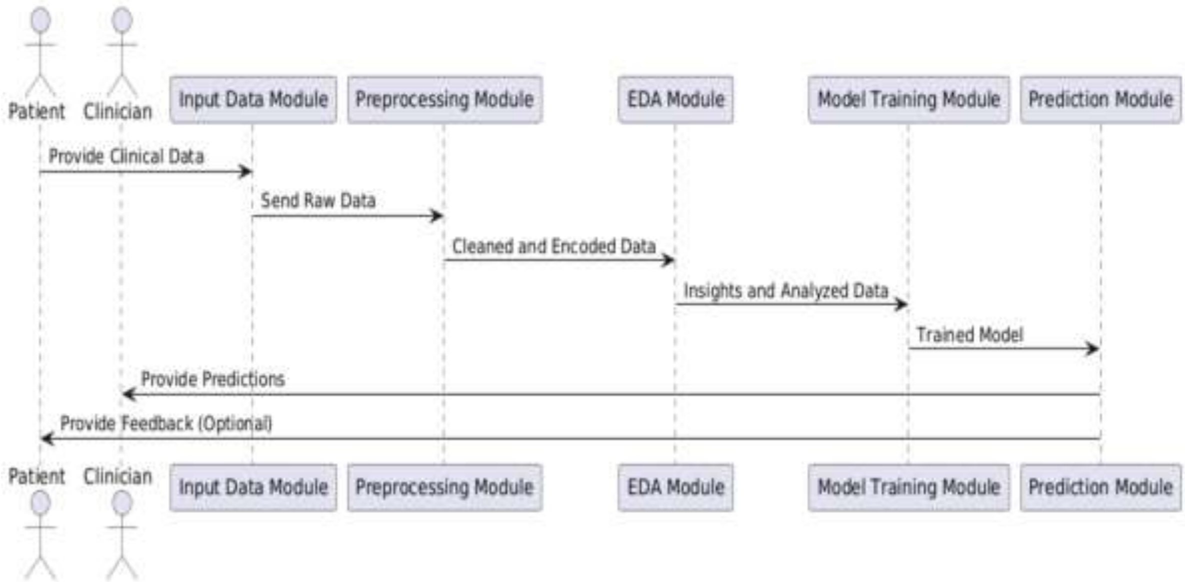

Fig 3.1.4 Sequence Diagram

3.1.5 ACTIVITY DIAGRAM

The activity diagram Fig 3.2 shows the workflow of a comprehensive healthcare data pipeline integrated with machine learning models for improved clinical decision-making. The process begins with Patient Registration, where patients are entered into the system. Next, Clinical Data Collection gathers relevant health information, including patient history, test results, and other clinical metrics. This data is then consolidated during the Data Integration phase, where it is formatted and stored within the system. Once integrated, the data undergoes Data Preprocessing and Cleaning to remove inconsistencies, handle missing values, and ensure high-quality data for analysis. During the Feature Extraction and Selection phase, the most relevant features are identified to optimize the performance of machine learning models. The processed data is then fed into TabNet, a model designed for tabular data, followed by Vision Transformer Deployment, which excels in analyzing image-based or complex data. The outputs from these models are synthesized in the Result Generation phase, producing actionable insights or predictions. These results guide Clinical Actions, such as diagnoses, treatment planning, or monitoring. This pipeline ensures a streamlined, data-driven approach to healthcare.

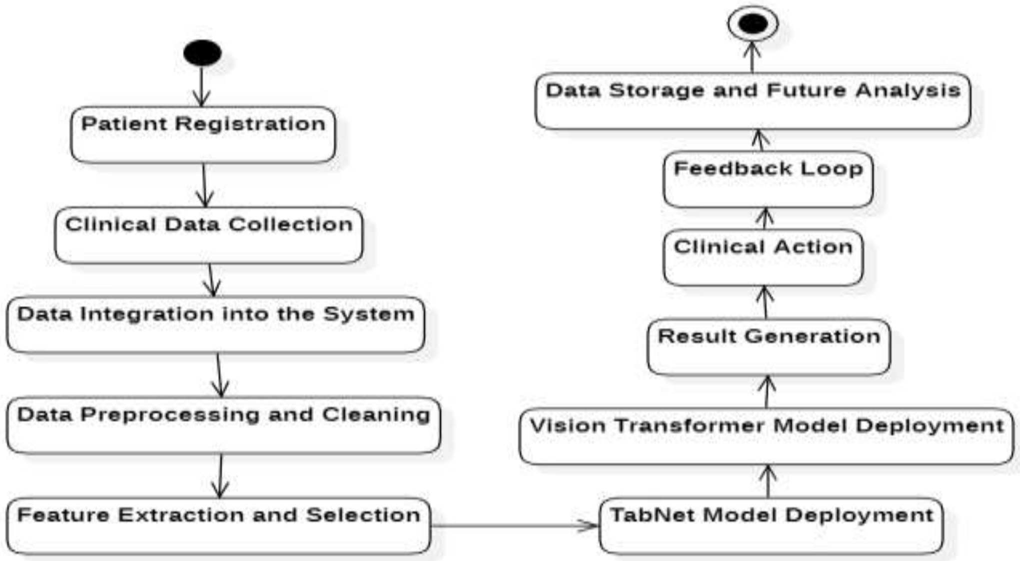

Fig 3.1.5 Activity Diagram

CHAPTER 4
PROJECT DESCRIPTION

4.1 MODULES

4.1.1 DATA PREPROCESSING:

The preprocessing module cleans and prepares the integrated data. It handles missing values through statistical imputation, normalizes continuous variables to reduce the impact of outliers, and encodes categorical features like Gender and Ethnicity using one-hot encoding. This step ensures a consistent dataset for modeling.

4.1.2 EXPLORATORY DATA ANALYSIS

The EDA module analyzes the dataset to uncover patterns and relationships. It examines the distributions of variables, identifies correlations among features, and detects potential outliers. Visualizations such as histograms, scatter plots, and heatmaps are generated to guide feature selection and model training.

4.1.3 TABNET MODEL TRAINING:

This module focuses on implementing the TabNet deep learning model. It processes the preprocessed data, dynamically selects relevant features using attention mechanisms, and trains the model to predict neurocognitive impairments. Hyperparameter tuning and cross-validation are included to ensure the model's robustness and generalizability.

4.1.4 PREDICTION AND EVALUATION:

This module generates predictions of neurocognitive disorder risk levels for individual patients. It evaluates the performance of the trained TabNet model using metrics like accuracy, precision, and F1-score. Feature importance scores are provided to help clinicians understand the key factors influencing each prediction, aiding in informed decision-making.

CHAPTER 5
IMPLEMENTATION

.

EDA

In the Exploratory Data Analysis (EDA) phase, the data underwent thorough cleaning to ensure its quality and readiness for analysis. Missing values were identified and addressed by omitting null entries or imputing them based on statistical methods such as mean for continuous variables and mode for categorical variables. Feature extraction played a crucial role in selecting the most relevant attributes. A correlation matrix was utilized to identify relationships between features, helping to isolate the ones with the highest predictive significance for neurocognitive disorders.

Several features were visualized using bar graphs, box plots, and scatter plots to analyze distributions and detect outliers. For instance, box plots were used to examine variables like BMI and blood pressure, revealing the presence of extreme values that were either adjusted or removed to improve data consistency. The insights from these visualizations guided the selection of important features, ensuring that only the most impactful variables, such as FamilyHistoryAlzheimers and MMSE scores, were retained for model training. This comprehensive EDA process helped streamline the dataset, eliminating irrelevant data and preparing it for optimal performance with the TabNet model.

EDA is crucial for this dataset as it lays the foundation for accurate predictions by ensuring the data is clean, consistent, and representative of the problem at hand. Neurocognitive disorders are influenced by a wide range of factors, including demographics, lifestyle habits, and medical history. Without a thorough understanding of the dataset, it would be challenging to capture the subtle relationships between these variables. For instance, EDA helps identify correlations between FamilyHistoryAlzheimers and MMSE scores, which are key predictors of cognitive decline. By visualizing features through box plots and bar graphs, we can detect outliers and imbalances that may skew model training. Additionally, reducing the dataset to the most relevant features improves model efficiency and interpretability. EDA not only prepares the dataset for the TabNet model but also ensures that the insights gained are grounded in meaningful data patterns, enhancing the system's reliability.

TABNET

The TabNet model implementation involved fine-tuning various hyperparameters to ensure optimal performance on the prepared dataset. The number of epochs was set to 150 to allow the model sufficient iterations to learn the patterns within the data. Other parameters, such as learning rate, batch size, and the sparsity coefficient, were carefully configured to balance training efficiency and accuracy. These hyperparameters were tuned using a grid search method to identify the best possible combination for this specific dataset.

During training, TabNet leveraged its dynamic feature selection mechanism, which allowed the model to focus on the most predictive features during each decision-making step. This ensured that important variables, like MMSE scores and FamilyHistory, were given higher priority while less relevant ones were downweighted. The model's ability to adaptively select features not only improved its interpretability but also enhanced its predictive accuracy. Cross-validation was employed to assess the model's robustness, ensuring it generalized well across different data subsets. After training, the model was evaluated on unseen data to validate its effectiveness in predicting neurocognitive disorders.

TabNet's unique capabilities make it particularly suited for this dataset, where understanding the importance of each feature is critical for both accuracy and interpretability. Unlike traditional machine learning models, TabNet's dynamic feature selection mechanism allows it to prioritize key indicators such as MMSE scores and FamilyHistory, which have a significant impact on the diagnosis of neurocognitive disorders. This is vital in medical datasets where certain features inherently carry more weight than others. Furthermore, TabNet's interpretability ensures that the model's predictions are transparent, enabling clinicians to trust the results and understand the factors driving them. By tuning hyperparameters and leveraging its sequential attention mechanism, TabNet optimally processes the complex interactions between features in this dataset, providing precise and actionable predictions that directly support early diagnosis and effective intervention strategies.

5.1 OUTPUT SCREENSHOTS

On observing the dataset using bar charts, with the features: Smoking, CardioVascularDisease, Diabetes, HeadInjury, MemoryComplaints and FamilyHistory against count of diagnosed individuals, it was observed that family history and smoking played a major role in prognosis of neurocognitive impairment. Memory complaints was also prominently faced by diagnosed individuals. Although head injury is the highest on the scale, it can be seen that head injury does not account or play a major role in getting any form of dementia.

On further analyzing the features, this study visualized 18 key features using a heatmap. Memory complaints and behavioral problems were highly correlated with diagnosis which showed that these were the highest popular symptoms among the individuals with neurocognitive impairment. It can also be seen that the correlation among the other features were comparatively lower, which also shows why medical professionals have a hard time detecting these diseases. The implementation of deep learning algorithms in this field would definitely prove to be fruitful.

With medians of roughly 6 and 5, respectively, Physical Activity and Diet Quality show modest average scores but notable variability, with a wide range of 1 to 10. The range of sleep quality scores is from 2 to 10, with a higher median score near 7 indicating that patients have better average sleep quality. Activities of Daily Living (ADL) and Functional Assessment scores, on the other hand, are lower, with ranges between 1 and 8 and medians close to 3 and 4, respectively, suggesting that these areas may be more difficult for people with this diagnosis.As it can be seen the TabNet algorithm’s confusion matrix, true positive and true negative’s count is higher than the falsely classified count, which shows that the algorithm is performing really well.

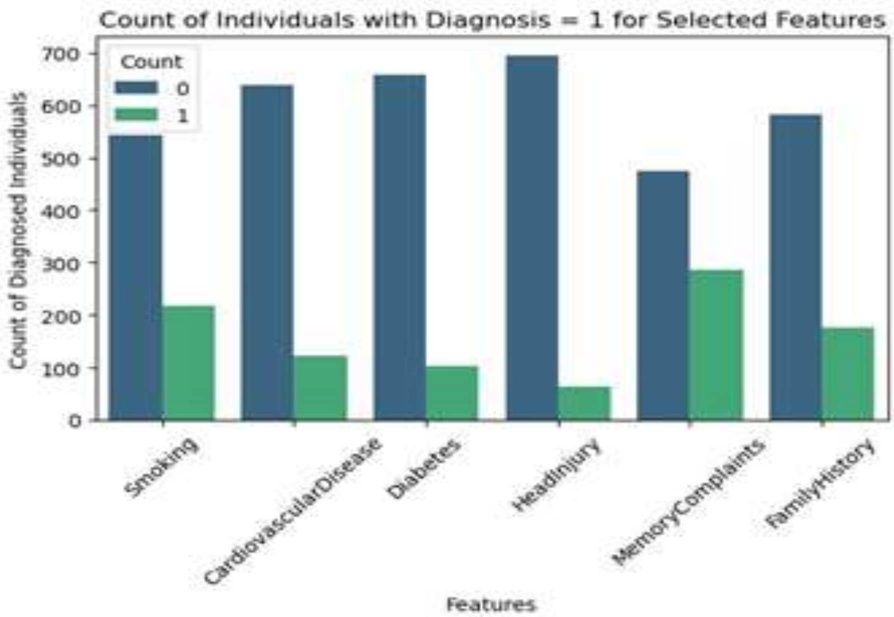

Figure 5.3.1 Key Features vs Diagnosed Count

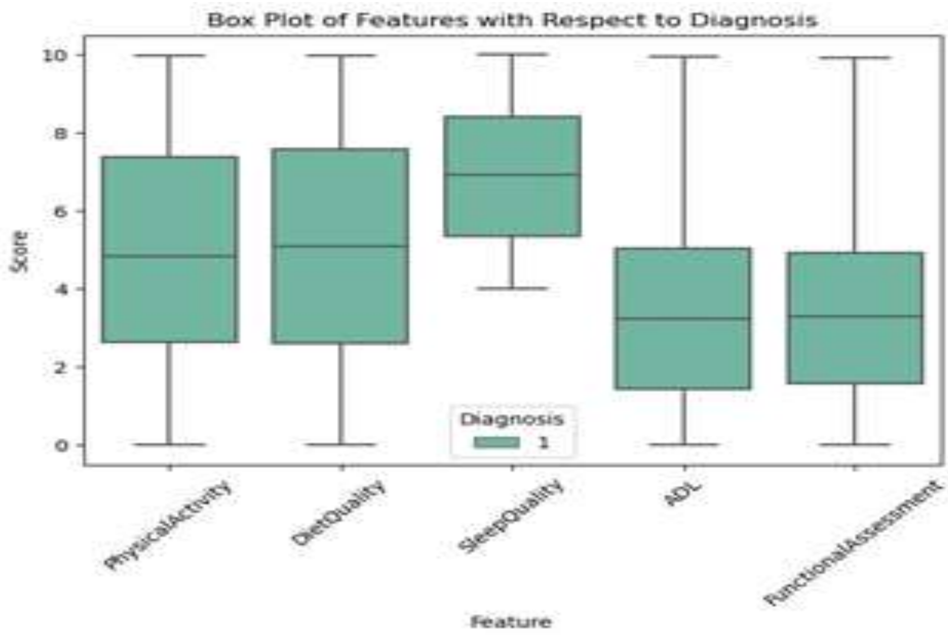

Figure 5.3.2 Box-Plot

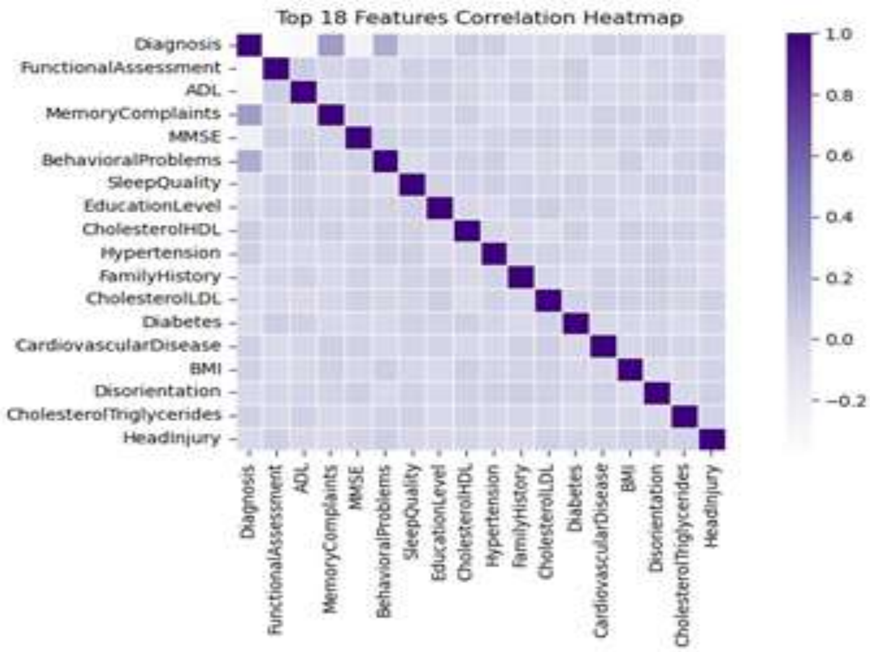

Figure 5.3.3 Correlation among 18 features

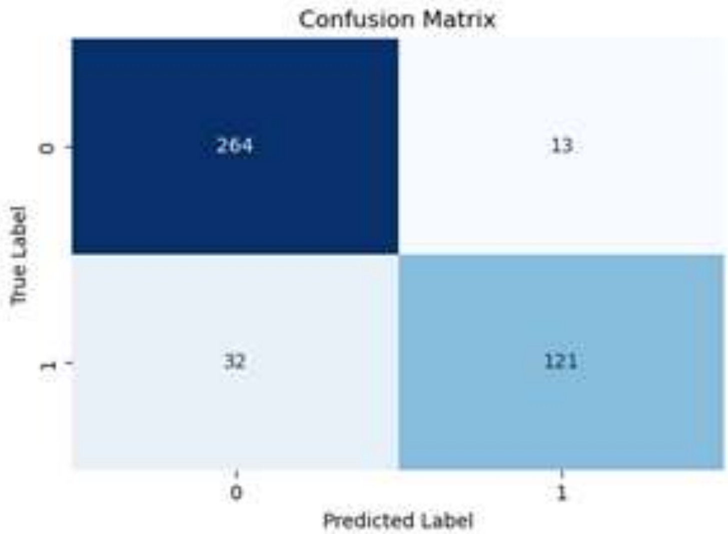

Figure 5.3.4 Confusion Matrix

Figure 5.3.5 Training loss over epochs

Precision	Recall	F1-score	Accuracy
0.89	0.82	0.85	0.90

Figure 5.3.6 Performance metrics for TabNet

CHAPTER 6

CONCLUSION AND WORK SCHEDULE FOR PHASE II

6.1 CONCLUSION

The development of a deep learning-based system for early detection of neurocognitive disorders represents a significant step forward in healthcare, particularly for elderly populations at risk of diseases like Alzheimer's and dementia. Traditional diagnostic methods often rely on cognitive assessments and imaging techniques, which, while effective, are costly, time-consuming, and typically used when symptoms have already advanced. This project aims to address these limitations by leveraging readily available clinical data to predict neurocognitive impairments at an early stage, allowing for timely interventions and better management of patient outcomes.

Through the implementation of the TabNet model, which excels at handling structured clinical data, this system offers a solution that dynamically prioritizes the most relevant features for accurate predictions. By utilizing a combination of demographic information, medical history, lifestyle factors, and cognitive assessments, the system makes it possible to predict the likelihood of developing a neurocognitive disorder with high accuracy. Furthermore, the system's ability to identify the most significant contributing factors, such as family history or specific cognitive symptoms, provides valuable insights that can guide clinicians in decision-making and treatment planning.

In conclusion, this project highlights the potential of artificial intelligence in transforming neurocognitive disorder diagnostics. By combining routine clinical data with powerful deep learning techniques, the system not only improves the accuracy and accessibility of early diagnosis but also offers a framework for future enhancements that can continue to evolve as new data and technologies emerge. The project provides a foundation for integrating AI into everyday healthcare practices, making early detection of neurocognitive disorders more efficient, accessible, and impactful.

6.2 WORK SCHEDULE FOR PHASE II

The future version of this model will integrate a Vision Transformer to analyze image data alongside tabular data. This multi-modal approach will enhance detection methods, leading to more accurate diagnoses. By incorporating images of demented brains, the model aims to offer a deeper understanding of neurocognitive disorders. This combination of data types will allow for better identification of cognitive decline at various stages. The goal is to improve early detection and intervention, helping patients receive care before crossing critical thresholds. By leveraging both imaging and clinical data, the model will provide a comprehensive diagnostic tool. This approach will contribute to more effective treatment strategies. The integration of these technologies is expected to strengthen the overall accuracy of neurocognitive disorder detection. Ultimately, this will help clinicians intervene earlier, improving patient outcomes. Through this multi-modality, neurocognitive disorders can be better understood and managed.

REFERENCES

[1] Heising, L. M., & Angelopoulos, S. (2021). “Early Diagnosis of Mild Cognitive Impairment with 2-Dimensional Convolutional Neural Network Classification of Magnetic Resonance Images”. In Proceedings of the 54th Hawaii International Conference on System Sciences.

[2] Pacheco-Lorenzo, M. R., Valladares-Rodríguez, S. M., Anido-Rifón, L. E., & Fernández-Iglesias, M. J. “Smart Conversational Agents for the Detection of Neuropsychiatric Disorders: A Systematic Review”. Department of Telematics Engineering at University of Vigo, Escola de Enxeñaría de Telecomunicación, Campus Lagoas-Marcosende, 36310 Vigo, Spain.

[3] P. Kumar, V. Subathra, Y. Swasthika and V. Vishal, "Disease Diagnosis Using Machine Learning on Electronic Health Records," 2024 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2024, pp. 1-6, doi: 10.1109/IC3IoT60841.2024.10550235.

[4] Shah, S. M. A. H., Khan, M. Q., Rizwan, A., Jan, S. U., Abdel Samee, N., & Jamjoom, M. M. “Computer-Aided Diagnosis of Alzheimer’s Disease and Neurocognitive Disorders with Multimodal Bi-Vision Transformer (BiViT”). Received: 2 August 2023 / Accepted: 14 June 2024.

[5] Sirilertmekasakul, C., Rattanawong, W., Gongvatana, A., & Srikiatkhachorn, A. “The Current State of Artificial Intelligence-Augmented Digitized Neurocognitive Screening Test”. Received: 29 December 2022 / Accepted: 20 March 2023 / Published: 30 March 2023.

[6] Ghoraani, B., Boettcher, L. N., Tolea, M. I., Galvin, J. E., Hssayeni, M. D., & Rosenfeld, A. “Detection of Mild Cognitive Impairment and Alzheimer’s Disease Using Dual-Task Gait Assessments and Machine Learning”. Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, 33431, USA / Comprehensive Center for Brain Health, Department of

Neurology, University of Miami, Miami, FL, 33136, USA.

[7] Rutkowski, T. M., Abe, M. S., Koculak, M., & Otake-Matsuura, M. “Classifying Mild Cognitive Impairment from Behavioral Responses in Emotional Arousal and Valence Evaluation Task–AI Approach for Early Dementia Biomarker in Aging Societies”.

[8] Javeed, A., Dallora, A. L., Rauf, H. T., Berglund, J. S., & Anderberg, P. “Early Prediction of Dementia Using Feature Extraction Battery (FEB) and Optimized Support Vector Machine (SVM) for Classification”. Received: 15 January 2023 / Revised: 30 January 2023 / Accepted: 31 January 2023 / Published: 2 February 2023.

[9] Kamal, M. S., Northcote, A., González Crespo, R., Chowdhury, L., Dey, N., & Herrera-Viedma, E. “Alzheimer’s Patient Analysis Using Image and Gene Expression Data and Explainable-AI to Present Associated Genes”. IEEE Transactions on Instrumentation and Measurement, VOL. 70, 2021, 2513107.

[10] Tsiakiri, A., Bakirtzis, C., Plakias, S., Terzoudi, A., & Christidi, F. “Predictive Models for the Transition from Mild Neurocognitive Disorder to Major Neurocognitive Disorder: Insights from Clinical, Demographic, and Neuropsychological Data”. Received: 27 April 2024 / Revised: 27 May 2024 / Accepted: 29 May 2024 / Published: 1 June 2024.

[11] R. K. Mahendran, R. Aishwarya, R. Abinayapriya and P. Kumar, "Deep Transfer Learning Based Diagnosis of Multiple Neurodegenerative Disorders," 2024 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2024, pp. 1-4, doi: 10.1109/ESCI59607.2024.10497320.

[12] AlHarkan, K., Sultana, N., Al Mulhim, N., AlAbdulKader, A. M., Alsafwani, N., Barnawi, M., Alasqah, K., Bazuhair, A., Alhalwah, Z., Bokhamseen, D., Aljameel, S. S., Alamri, S., Alqurashi, Y., & Al Ghamdi, K.” Artificial Intelligence Approaches for Early Detection of Neurocognitive Disorders Among Older Adults”. Received: 04 October 2023 / Accepted: 29 January 2024 / Published: 16 February 2024.

[13] Lagergren, M., Fratiglioni, L., Rahm Hallberg, I., Berglund, J., Elmståhl, S., Hagberg, B., Holst, G., Rennemark, M., Sjölund, B.-M., Thorslund, M., Wiberg, I., Winblad, B., & Wimo, A. “A Longitudinal Study Integrating Population, Care and Social Services Data: The Swedish National Study on Aging and Care (SNAC)”.

[14] Luz, S. “Longitudinal Monitoring and Detection of Alzheimer’s Type Dementia from Spontaneous Speech Data”. In 2017 IEEE 30th International Symposium on Computer-Based Medical Systems.

[15] Islam, R., & Layek, M. A. “StackEnsembleMind: Enhancing Well-Being Through Accurate Identification of Human Mental States Using Stack-Based Ensemble Machine Learning”. Department of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh, 2023.

[16] A. Sivaprakasam S, S. P. S, V. SP, V. M and S. J, "A Novel Approach on Parkinson Detection with Deep Ensemble Network," 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Krishnankoil, Virudhunagar district, Tamil Nadu, India, 2024, pp. 1-5, doi: 10.1109/INCOS59338.2024.10527508.

APPENDIX

APPENDIX 1:
LIST OF PUBLICATIONS

PUBLICATION STATUS: Submitted to the conference.

TITLE OF THE PAPER: Detection Of Neurocognitive Impairment In The Elderly Using TabNet Algorithm

AUTHORS: Ananthajothi K, Oviya K, Navya Balasundaram

NAME OF THE CONFERENCE: International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES -2024), St. Joseph's Institute of Technology.

DATE OF PRESENTATION: 12th or 13th December 2024

ICSES -2024 Conference –Decision – Reg.

1 message

ICSES St. Joseph's Institute of Technology <icses@stjosephstechnology.ac.in> Wed, 13 Nov, 2024 at 3:38 pm

To: ananthajothi.k@rajalakshmi.edu.in, 210701186@rajalakshmi.edu.in, 210701177@rajalakshmi.edu.in

Dear Author(s),

On behalf of the conference committee, we are happy to inform you that your paper ID: ICSES-24-T3-551 and title "Detection Of Neurocognitive Impairment In The Elderly Using TabNet Algorithm" has been accepted.

We invite you to present your research paper in **physical mode** at the International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES -2024).

Registration Information is available at https://icses2024.stjosephstechnology.ac.in/register.php

After the payment, kindly send the final paper and proof of payment to icses@stjosephstechnology.ac.in on or before 15th November 2024.

Conference Date: 12th & 13th December, 2024.

While sending your manuscript, mention your paper id in the subject line and rename your file name as paper id without fail.

Best regards,

Organizing Committee,

ICSES -2024,

St. Joseph's Institute of Technology, OMR, Chennai-119.

icses@stjosephstechnology.ac.in

Whatsapp : +91 99443 37181

APPENDIX 2:

SAMPLE CODE:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv("C:\\Users\\OVIYA\\OneDrive\\Desktop\\final yr
proj\\datasets\\alzheimers_disease_data.csv")
df.rename(columns={'FamilyHistoryAlzheimers': 'FamilyHistory'}, inplace=True)
# Assuming your DataFrame is 'df' and has been loaded with top features
# Features and target column selection
features = [
    'FunctionalAssessment', 'ADL', 'MemoryComplaints', 'MMSE', 'BehavioralProblems',
    'SleepQuality', 'EducationLevel', 'CholesterolHDL', 'Hypertension',
    'FamilyHistory', 'CholesterolLDL', 'Diabetes', 'CardiovascularDisease',
    'BMI', 'Disorientation', 'CholesterolTriglycerides', 'HeadInjury'
]
target = 'Diagnosis'

X = df[features].values
y = df[target].values
# List of columns to exclude
columns_to_exclude = ['Gender', 'Ethnicity', 'EducationLevel', 'Smoking',
'FamilyHistoryAlzheimers', 'CardiovascularDisease', 'Diabetes',
        'Depression', 'HeadInjury', 'Hypertension', 'MemoryComplaints',
'BehavioralProblems', 'Confusion', 'Disorientation',
        'PersonalityChanges', 'DifficultyCompletingTasks', 'Forgetfulness', 'Diagnosis']

# Select the rest of the columns using list comprehension
num_columns = [col for col in data1.columns if col not in columns_to_exclude]
import matplotlib.pyplot as plt
```

```
# Plot histograms for numeric columns
data1[num_columns].hist(bins=15, figsize=(15, 10))
plt.show()
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Initialize the TabNet classifier
clf = TabNetClassifier()

# Train the model
clf.fit(
    X_train, y_train,
    max_epochs=150,
    patience=20,
    batch_size=1024,
    virtual_batch_size=128,
    num_workers=0,
    drop_last=False
)

# Predict and evaluate
y_pred = clf.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"Model accuracy: {accuracy * 100:.2f}%")

# Get the correlations with respect to the chosen target column
target_corr = corr_matrix['Diagnosis']

# Step 3: Sort the correlations by absolute value and select the top 18 features
top_features = target_corr.abs().sort_values(ascending=False).head(18)

# Display the top features
print("Top 18 features correlated with", 'Diagnosis')
```

```
print(top_features)

# Optional: Visualize these top features in a heatmap
plt.figure(figsize=(10, 6))
corr_matrix.values[np.diag_indices_from(corr_matrix)] *= 0.5
sns.heatmap(df[top_features.index].corr(), annot=True, fmt=".2f", cmap='Purples',
            linewidths=.5, cbar=True, square=True)

# Set the title and labels
plt.title(f'Top 18 Features Correlation Heatmap')
plt.xticks(rotation=90)
plt.yticks(rotation=0)
plt.tight_layout()

# Show the plot
plt.show()

from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay
import seaborn as sns
import matplotlib.pyplot as plt

# Predict on the test set
y_pred = clf.predict(X_test)

# Generate confusion matrix
cm = confusion_matrix(y_test, y_pred)

# Visualize the confusion matrix with heatmap
plt.figure(figsize=(6, 4))
sns.heatmap(cm, annot=True, fmt="d", cmap="Blues", cbar=False)
plt.xlabel("Predicted Label")
plt.ylabel("True Label")
plt.title("Confusion Matrix")
```

```
plt.show()
# Predict on the test set
y_pred = clf.predict(X_test)

# Calculate Precision, Recall, and F1 Score
precision = precision_score(y_test, y_pred, average='binary')
recall = recall_score(y_test, y_pred, average='binary')
f1 = f1_score(y_test, y_pred, average='binary')
print(f"Model accuracy: {accuracy * 100:.2f}%")
print(f"Precision: {precision:.2f}")
print(f"Recall: {recall:.2f}")
print(f"F1 Score: {f1:.2f}")



# Capture the losses from the history after training
if hasattr(clf, 'history'):
    epoch_losses = clf.history['loss']  # Extract the loss values directly

# Plotting the loss values against epochs
plt.figure(figsize=(8, 4))
plt.plot(range(1, max_epochs + 1), epoch_losses, linestyle='-', color='r', label='Training
Loss')
plt.title('Training Loss Over Epochs')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.xticks(range(1, max_epochs + 1, 10))  # Ticks every 10 epochs
plt.grid()
plt.legend()
plt.tight_layout()
plt.show()
```

RE-2022-424641-plag-report

ORIGINALITY REPORT

3%	2%	1%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	1%
2	Submitted to Colorado Technical University Online Student Paper	<1%
3	www.devx.com Internet Source	<1%
4	researchspace.ukzn.ac.za Internet Source	<1%
5	erules.veristar.com Internet Source	<1%
6	worldwidescience.org Internet Source	<1%
7	www.researchgate.net Internet Source	<1%
8	eprints.kfupm.edu.sa Internet Source	<1%

Detection Of Neurocognitive Impairment In The Elderly Using TabNet Algorithm

Dr Ananthajothi K,
Department of Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
ananthajothi.k@rajalakshmi.edu.in

Oviya K,
Department of Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
210701186@rajalakshmi.edu.in

Navya Balasundaram,
Department of Computer Science and Engineering
Rajalakshmi Engineering College
Chennai, India
210701177@rajalakshmi.edu.in

***Abstract*—As the global population ages, neurocognitive disorders like the Alzheimer's disease which is a form of dementia, and several other forms of dementia are becoming more commonplace. These conditions significantly impair the quality of life for patients and present a growing challenge for healthcare systems worldwide. Early detection is critical to improving patient outcomes, yet traditional diagnostic methods often struggle to catch on to the very early and subtle signs of decline in cognitive function. In this paper, a deep learning algorithm, TabNet, is used on clinical data to detect neurocognitive impairment. Before applying the algorithm, Exploratory Data Analysis is also performed to get a better understanding of the dataset. This research is mainly focused on people with family history. The TabNet model, designed especially for tabular data easily and effectively learns the importance of each feature and their inter-dependencies. In future work, it is aimed to make this research more holistic by using Vision Transformer models to analyze MRI images. This study currently focuses on clinical data, which includes demographic, lifestyle, and cognitive assessments, and demonstrates how TabNet can be used to identify signs of cognitive impairment. Working with critical features, this study also focuses on data pre-processing and feature selection. This project's main aim is to utilize AI in healthcare to help with diagnoses that are otherwise somewhat difficult to make.**

Keywords- Neurocognitive Disorders, Alzheimer's Disease, TabNet, Exploratory Data Analysis(EDA), Clinical Data, Feature Selection, Deep Learning, Elderly Population, Early Diagnosis, Prediction Model, AI in Healthcare

I. Introduction

Neurocognitive disorders like Alzheimer's, Parkinson's, and several varieties of dementia are increasingly common as life expectancy rises across the globe. These conditions mostly begin with subtle cognitive symptoms that are difficult to detect using traditional diagnostic methods. Unfortunately, substantial brain damage might have already taken place by the time these symptoms are sufficiently evident for standard evaluation tools to detect them. Because of this, patients are frequently diagnosed later in life, when there are fewer options for treatment, which lowers the probability of an effective intervention.

Given how crucial early detection is, artificial intelligence (AI) presents an acceptable approach. Large volumes of clinical data can be analyzed by AI, particularly deep learning, to find hidden patterns that might not be immediately obvious to human observers. In the area of detecting neurocognitive disorders, where immediate diagnosis is necessary to improving patient outcomes, this ability is particularly helpful. Additionally, early diagnosis can help with interventions that improve quality of life, lessen symptom severity, and slow the progression of the disease.

In this work, the concentration is on clinical data from elderly patients to predict neurocognitive disorders using the TabNet deep learning algorithm. Since TabNet can handle tabular datasets and dynamically learn which features are most important for prediction, it will be suitable for this study. The goal is to use this model to create a diagnostic tool that will enable medical professionals to detect cognitive decline earlier and with greater accuracy than is possible with conventional techniques. Future research will include MRI images analyzed with the Vision Transformer model to offer a more thorough picture of brain health. This multimodal strategy will combine behavioral, cognitive, and structural data to further improve diagnostic accuracy. Additionally, understanding the dataset prior to any implementation is crucial. So, Exploratory Data Analysis is performed to understand the features better, which will be more insightful along with TabNet results.

The clinical dataset used in this study includes demographic information, lifestyle factors, medical history, and cognitive assessments, all of which are potential indicators of neurocognitive disorders. By using these input features, the TabNet model is trained to recognize the early stages of cognitive impairment with high accuracy. It will learn from every single feature and understand them thoroughly, it will understand their dependencies and use that to make the model more accurate. The integration of imaging data will be explored in future research and be combined with these insights gained from clinical assessments, thus providing a more holistic prediction system.

II. Related Works

This paper systematically reviews the use of SCAs [2] for neuropsychiatric disorder detection. It identified 17 studies based on disorders like depression, dementia, and other impairments in cognition, by using AI technologies in feature extraction and analysis through deep learning methods. Most systems that are related to verbal and audiovisual data for prediction validation lack comprehensiveness, especially as regards reliability and usability. Though the field holds great promise and is increasingly integrating AI, current SCAs are not yet validated for clinical use.

Predictive models were created in a different study [10] to evaluate the change from mild to serious neurocognitive dysfunction. Over a two-year period, 132 subjects' clinical, demographic, and neuropsychological data were incorporated into the model. It identified important risk variables such as high body mass index and alcohol intake with an accuracy of 83.7%. The danger was decreased by characteristics like being female and having better praxis ability. The generalizability of the model was constrained by the small dataset.

By combining a Support Vector Machine (SVM) with a Feature Extraction Battery (FEB) [8], the proposed FEB-SVM algorithm sought to enhance dementia prediction. The model outperformed 12 state-of-the-art techniques, reaching accuracy of 98.28% on training set and accuracy of 93.92% on testing set. FEB's created features were successful in increasing prediction accuracy, but they were ineffective in pinpointing the precise causes of dementia. More feature development was stressed throughout the study in order to gain deeper insights.

Population and care data were merged in a Swedish longitudinal study (SNAC) [13] to record the aging process and care requirements in various locations. Many aspects of senior health, care, and social services were examined in this study. The multidisciplinary methodology made it possible to analyze aging holistically. Unfortunately, the extensive breadth necessitated complicated and resource-intensive data collection and study design. The article emphasized the difficulties of overseeing a dataset this size.

In a study of Saudi patients, an AI-based method [12] for diagnosing neurocognitive problems that employed logistic regression and SVM models achieved up to 95.5% accuracy. Important variables like chronic illnesses and educational attainment were brought to light by the study. Even though the SVM model worked well, more research was needed to determine whether it could be applied to other populations. The study found that in order to improve generalizability, it is crucial to use a variety of datasets.

Using 2D MRI data, a Bi-Vision Transformer (BiViT) model [4] provided a novel transformer-based method for classifying cognitive diseases, including Alzheimer's disease. The model used cutting-edge feature learning modules to achieve 96.38% accuracy. It captured intricate patterns in medical photos better than previous deep learning techniques. Its application in real-world scenarios is constrained in the absence of more data due to its poor performance in smaller or unbalanced datasets.

An analysis of AI-augmented neurocognitive screening tests [5] examined the ways in which deep learning and speech recognition are examples of AI and machine learning techniques that have enhanced cognitive evaluations. The study covered how AI might improve exam accuracy and lessen biases. Notwithstanding these developments, issues with technological accessibility, data privacy, and dependability nonetheless exist. It was noted that older-friendly designs and more validation were required.

By merging weak learner models, the StackEnsembleMind [15] model improved mental health assessments through a stack-based ensemble machine learning technique. The model identified mental states including stress and worry with 98% accuracy. Reliability was increased by addressing class imbalances with SMOTE. Unfortunately, the model's intricacy limited its practical application by making it challenging to understand and apply in healthcare settings.

Another paper explores the application of machine learning, particularly deep learning [11] , for the classification of three neurodegenerative diseases: cerebral ataxia, Alzheimer's disease, and Parkinson's disease. The authors uses a two-step methodology, employing pre-trained neural network architectures in the form of VGG16 and ResNet101. In fine-tuning these models for each disease, the researchers enhance the networks' ability to differentiate among them. Optimizations techniques like SGD, Rmsprop, and Adam are used to get the best result; evaluation metrics such as accuracy, loss are applied for the assessment of the model. The model is deployed through the Flask web framework, making it accessible and user-friendly.

A paper discusses the use of machine learning algorithms in disease diagnosis, focusing on the application of these algorithms in Electronic Health Records (EHRs) [3]. Because of the ever-expanding patient data, machine learning is promising for improving the efficiency and accuracy of diagnosis. The study uses anonymized EHRs to include the medical history, laboratory results, and clinical notes to build models for predicting diseases. Various ML algorithms were implemented and evaluated against their interpretability, computational efficiency, and diagnostic accuracy. The results show that machine learning-based methods are superior compared to conventional diagnostic methods, providing improved accuracy and efficiency in disease diagnosis and prediction.

In order to predict mild cognitive impairment (MCI) [7] based on emotional arousal and valence tasks, a study that classified MCI from behavioral responses using machine learning. The model's accuracy in differentiating between MCI and normal cognition was around 90%. Comparing this non-invasive technology to traditional ones revealed practical advantages. To capture more sophisticated signs of cognitive impairment, the fundamental components of the activities needed to be further refined.

Analysis of vocal characteristics [14] such as speech rate and voice activity in natural conversations was the main goal of a long duration study about the use of voice data for Alzheimer's disorder detection. The technique used a Bayesian classifier and obtained 68% accuracy. Non-invasive speech data was a promising early diagnosis technique because of how easy it was to use. The quantity of the dataset, however, hampered its accuracy, indicating the need for further data to enhance performance.

A machine learning-based dual-task gait assessment model [6] was developed to detect Alzheimer's disorder and mild to medium cognitive decline. When the DSM-5 criteria were compared to those of other diagnostic systems, the model discovered that DSM-5 detected more cases than DSM-IV. Still, diagnosis accuracy for less severe illnesses like MCI was only moderately high. Although more work needed to be done, the study indicated that the DSM-5 criteria could help with early cognitive decline identification.

MRI data was utilized to build a 2D Convolutional Neural Network (CNN) [1] to identify moderate cognitive impairment (MCI) at early stage, which produced a user-friendly diagnostic tool. The model showed promise for application in resource-constrained clinical settings and performed well even with little datasets. But because of system constraints, crucial preprocessing stages like skull stripping were omitted, which might have affected performance. The study made recommendations for future developments in data preparation.

Alzheimer's disease was diagnosed using a multimodal technique that included gene expression data and MRI images, using deep learning models like CNN and SpinalNet [9] A complete method of diagnosis was made possible by the combination of genetic and imaging data. However, the method's processing cost restricted its full capacity areas with insufficient resources. The study underlined the need for more effective implementations while also highlighting the possibilities of multimodal techniques.

A paper proposed a system that aims to detect Parkinson's Disease (PD) [16] at an early stage using Magnetic Resonance Imaging (MRI) and artificial intelligence. MRI is very essential in identifying the characteristics of Parkinson's features, and this study makes use of deep learning, more importantly, a Deep Ensemble Network, to enhance its detection process. The combined system uses a CNN model and LSTM optimized using the ADAM algorithm. The accuracies, precisions, recalls, and F1 values obtained are used to depict the performance of the proposed model, which attains an astonishing 98% accuracy. Comparison with existing approaches proves that the proposed method ensures effective early detection of PD.

III. PROPOSED METHODOLOGY

A. Problem Definition:

This paper aims to develop a predictive model to identify neurocognitive disorders based on clinical data from elderly patients. The scope of such data would include demographics, lifestyle, and cognitive assessment. Though the future work will be deployed with regard to integrating Vision Transformers for MRI image analysis into more proper diagnosis, this paper concentrates on the clinical data analysis via the use of the TabNet model.

B. Dataset:

The data set includes measures of many types of features that are known to be associated with neurocognitive degradation. These include demographics, such as Age and Gender, as well as Ethnicity. Lifestyle factors, such as Smoking, Alcohol_Intake, Phy_Activity, EatingQuality; and medical history items including Family_History, Diabetes, CardiovascularDisease are also included. Cognitive and behavioral assessment tools such as the Mini-Mental State Examination (MMSE) and markers of memory complaints, confusion, and disorientation are also included.

Dataset Description

The dataset's features are as follows:
Demographics: P_ID, Age, Gender, Ethnicity, Level_of_Education

Health Factors: BMI, Smoke_habits, Alcohol_Intake, Phy_Activity, Eating_Quality, Sleeping_Quality, Family_History, Cardiovascular_Disease, Diabetes, Depression, Head_Injuries, Hypertension, Systole_BP, Diastole_BP, Cholest_Total, Cholest_lLDL, Cholest_HDL, Cholest_Triglycerides

Cognitive/Behavioral Assessments: MMSE_Score, Function_Assessment, Memory_related_Complaints, Behavior_Problems, ADL, Confusion, Dis-orientation, Personality_Change, Difficulty_Completion_Tasks, Forgetfulness, Diagnosis

C. Data Preprocessing and Feature Selection

Exploratory Data Analysis is performed. It is used to summarize the dataset after preprocessing it. It is important for understanding the dataset and getting key insights. Preprocessing of data can be considered one of the most important steps of this study since it ensures the dataset is prepared enough to train the TabNet model. Missing values are addressed by imputing them using appropriate statistical methods for continuous variables through Mean, and for categorical, Mode. Other categorical variables like Gender and Ethnicity are encoded in one-hot encoding so that the format can change for the model to take. Continuous variables, such as BMI and blood pressure, are normalized in order to minimize the influence of outliers and skewness. Then univariate and bivariate analysis is done to get a proper comparison and understand the dependencies between the features. Univariate analysis gives the information about the distribution of each variable, while bivariate analysis gives information about the relation between each variable and the target variable. This study also analyzed the distribution of the target variable, to check if the dataset is imbalanced or balanced.

In addition, feature selection is another important part of the methodology. TabNet, unlike typical machine learning models, has an intrinsic feature selection mechanism- during training, it learns dynamically which features are most important. This is very valuable for the domain of medicine because there are certain health indicators that carry more predictive weight than others. Hence, for instance, the variables might be FamilyHistory or MMSE scores-these are much more indicative of cognitive decline than SleepQuality or AlcoholConsumption. In that way, TabNet is able to draw attention to such vital features, keeping it interpretable and predictions made therefrom on most relevant data.

D. System Architecture

Data preparation starts with the system architecture. The clinical dataset is cleaned and standardized. Missing values for continuous variables were filled by Mean. For categorical features such as Gender, missing values were filled using Mode, and one-hot encoding was applied to the categorical variables. Continuous features like BMI are normalized, which reduces the effect of outliers. All the data that is unnecessary was filtered out through feature engineering after preprocessing, leaving only the features useful in the prediction of neurocognitive disorders. This processes the data into the best optimum in respect of hyperparameters, which makes the TabNet model achieve better accuracy.

The process in EDA is to visualize how features are distributed to identify relationships with neurocognitive conditions. Sometimes, features like Age and Gender follow typical patterns because cognitive disorders are more common among elderly people. From histograms, preliminary sense of how these variables correlate with the outcome under consideration can be seen. For eg, Correlation analysis is utilized in mapping relations between factors like Hypertension and CardiovascularDisease. These are very closely associated factors. Last but not least, features like BMI are checked for outliers; clinically relevant values like high blood pressure are retained, but extreme values of BMI might get capped in the process to keep the model sharply focused. Box plots and Histograms are displayed as results of this EDA, to give the final key insights.

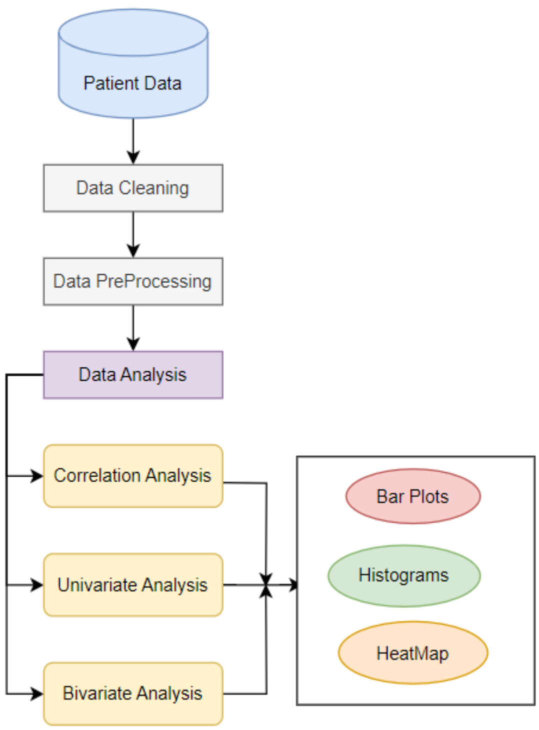

Figure 1: Working of EDA

In the neurocognitive disorder dataset, features have varying predictive strength, where some health indicators are more representative of a disease than others. The sequential attention mechanism in the model dynamically assigns greater attention to features at the decision step, such as FamilyHistory, MMSE scores, and Hypertension. The TabNet model will prefer the most informative features like this over the irrelevant ones, thereby filtering out the noise and focusing on high impact predictors. The fact that this model dynamically and adaptively focuses ensures that clinically important variables are emphasized for optimal accuracy and interpretability.

TabNet also has the ability to handle clinical data by transforming the raw data features with the feature transformer layers. These layers are applied to create more sophisticated representations, allowing the model to understand complex interactions among variables. For example, Age, SystolicBP, and CardiovascularDisease are feature interactions that influence the score of cognitive health all at once and are computed to highlight their interaction effect on neurocognitive disorders. The raw inputs being transformed to high-dimensional representations here enable deeper discovery of subtle dependencies relevant to making precise predictions.

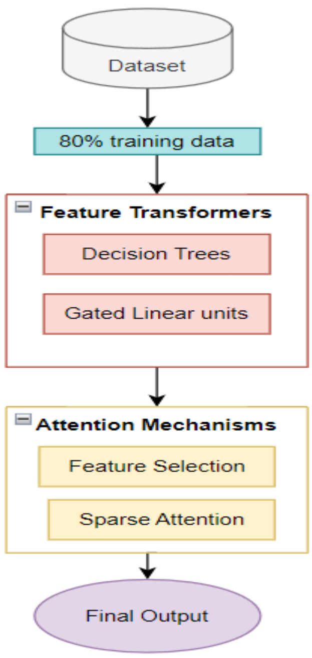

Figure 2: Working of TabNet Algorithm

Lastly, in the attentive transformer layers of TabNet, the refined feature representations are obtained through selective gating of only the most pertinent pathways. It utilizes the techniques of batch normalization and sparsity regularization to avoid overfitting and guarantee generalization to vastly diverse patient profiles in their health histories. This architecture improves interpretability through highlighting features most impactful on the predictions while making TabNet robust especially with complex clinical datasets. Following training, the model predicts neurocognitive impairment probability using a holistic view of the patient's health profile.

The system architecture thus begins with the use of EDA to clean and structure the clinical data in a clean manner so that it's ready for accurate analysis. After that, it is taken over by TabNet, which smartly pays attention to the specific factors like family history and cognitive scores in order to make precise predictions. Combining EDA groundwork with the focused learning of TabNet, this system effectively turns complex health data into meaningful predictions regarding support for the early identification of neurocognitive disorders. This model will prove to be even more useful for patients who had family history get themselves diagnosed, even when their symptoms are very subtle and early.

E. *Performance Metrics:*

Accuracy:

This metric evaluates the model’s overall prediction performance by determining the ratio of correct predictions—both true positives (TP) and true negatives (TN)—to the total instances in the dataset.

Precision:

Precision reflects the accuracy of the model’s positive predictions by calculating the proportion of true positives out of all instances labeled as positive. It is computed with the formula:

$$TP / (TP + FP)$$

where FP represents false positives and TP stand for True Positives.

Recall (Sensitivity):

Recall, or sensitivity, gauges the model’s effectiveness in identifying positive cases within the dataset. This metric is determined by the formula:

$$TP / (TP + FN)$$

where FN represents false negatives.

F1 Score: The F1 score combines both precision and recall into a single metric, offering a balanced view of the model's performance. It is the harmonic mean of these two metrics and is calculated as:

$$2 * (Precision * Recall) / (Precision + Recall).$$

Error checking

Confusion Matrix: This matrix organizes predictions into true positives, false positives, true negatives, and false negatives, allowing for a structured summary of the model's performance. From these values, various other metrics can be derived.

Results and Discussion

A. *Experimental Setup*:

The whole project was built on a Windows 11 operating system in a system with intel core i5 processor 3rd gen configuration, 8GB RAM, and 512GB SSD. The application is based on Python 3.11 and built using jupyter notebook from anaconda version 3. In Python, libraries: Panda, Seaborn and Matplotlib were used for Exploratory Data Analysis. PyTorch and sci-kit learn libraries were used additionally for TabNet algorithm. A lot of hyperparameter tuning was also done to achieve best results. Data was cleaned and pre-processed before usage in the experiment.

Observations:

On observing the dataset using bar charts, with the features: Smoking, CardioVascularDisease, Diabetes, HeadInjury, MemoryComplaints and FamilyHistory against count of diagnosed individuals, it was observed that family history and smoking played a major role in prognosis of neurocognitive impairment. Memory complaints was also prominently faced by diagnosed individuals. Although head injury is the highest on the scale, it can be seen that head injury does not account or play a major role in getting any form of dementia.

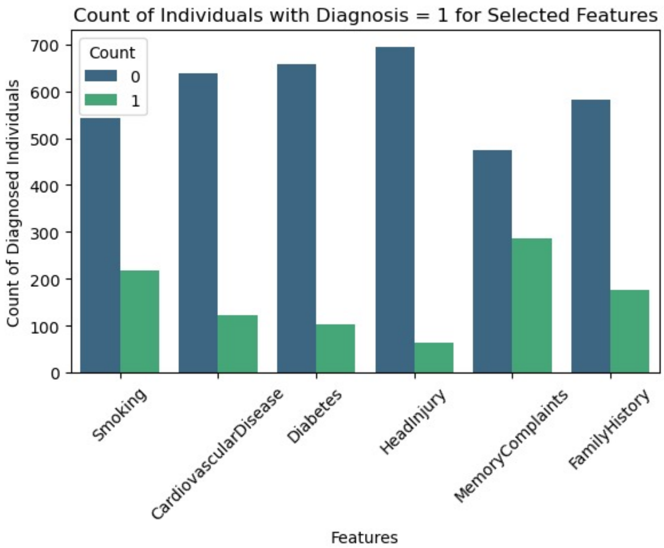

Figure 3: Key Features vs Diagnosed Count

On further analyzing the features, this study visualized 18 key features using a heatmap. Memory complaints and behavioral problems were highly correlated with diagnosis which showed that these were the highest popular symptoms among the individuals with neurocognitive impairment. It can also be seen that the correlation among the other features were comparatively lower, which also shows why medical professionals have a hard time detecting these diseases. The implementation of deep learning algorithms in this field would definitely prove to be fruitful.

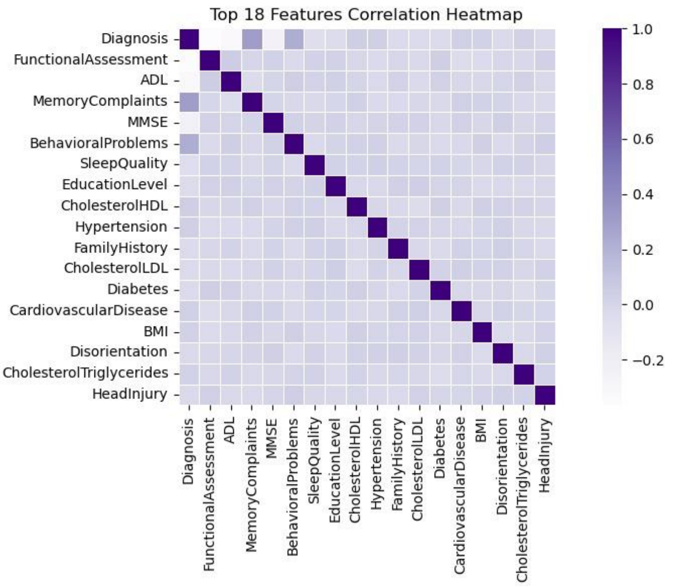

Figure 4: Correlation among 18 features

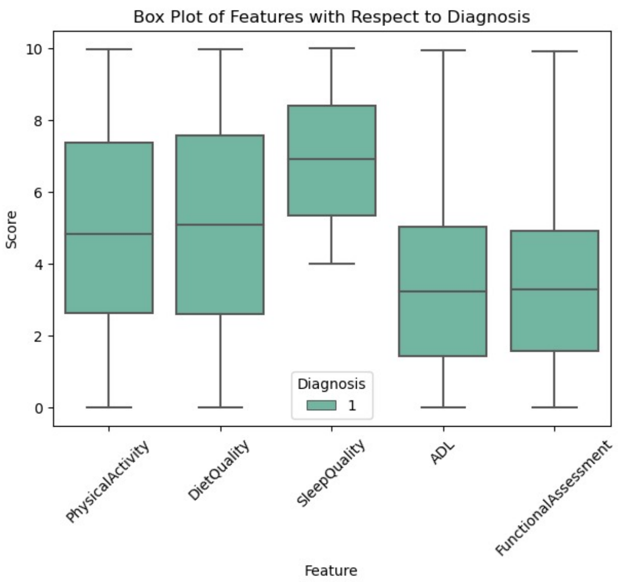

Figure 5: Box-Plot

Some other features were also visualized and analyzed using Box Plot, against a score with a scale of 0 to 10, with respect to diagnosis. Even though sleep quality seems to be comparatively better, ADL (Activities of Daily life such as consuming food, washing, writing, dressing-up etc) and Functional Assessments which also tests patient's ability to do simple tasks, has a very low score. These are most affected by neurocognitive impairment.

With medians of roughly 6 and 5, respectively, Physical Activity and Diet Quality show modest average scores but notable variability, with a wide range of 1 to 10. The range of sleep quality scores is from 2 to 10, with a higher median score near 7 indicating that patients have better average sleep quality. Activities of Daily Living (ADL) and Functional Assessment scores, on the other hand, are lower, with ranges between 1 and 8 and medians close to 3 and 4, respectively, suggesting that these areas may be more difficult for people with this diagnosis.

Precision	Recall	F1-score	Accuracy
0.89	0.82	0.85	0.90

Figure 6: Performance metrics for TabNet

Then TabNet algorithm is implemented on the dataset. For the hyperparameters, epochs count was kept at 150 and patience was kept at 20. Decision and attention dimension, unique to TabNet was kept at 16 and 16 respectively. Learning rate was 0.01. From experimental results, it can be seen that the accuracy is 90%. Precision is 0.89, Recall is 0.82 and F1 score is 0.85.

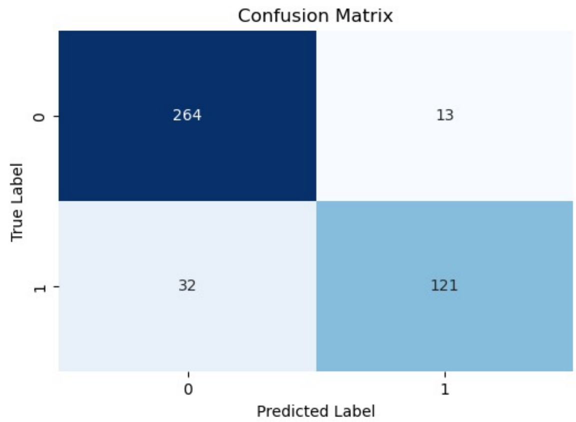

Figure 7: Confusion Matrix

A confusion matrix is a table that is easy to understand and helps to compare how well a classification model's predictions line up with the actual results. It categorizes predictions into four types: correct positives, correct negatives, and two kinds of mistakes: false positives and false negatives. This table helps to look at a glance where the model gets things right and where it goes wrong, making it easier to improve accuracy.

As it can be seen the TabNet algorithm's confusion matrix, true positive and true negative's count is higher than the falsely classified count, which shows that the algorithm is performing really well.

Figure 8: Training loss over epochs

This graph shows the training loss decreasing as it run more epochs. Training loss is usually high in the beginning as the model does not predict accurately, but as it passes more epochs, it learns the patterns and starts predicting with better accuracy, and hence the graph starts dipping. The curve starts at 0.9 during the 1st epoch and then starts falling and reaches below 0.2 at the 150th epoch. This means the algorithm is working well.

IV. Conclusion and Future Work

This study used the TabNet algorithm to analyze tabular data for the early identification of neurocognitive disorders. This model performed the function very well by providing an identification of complex patterns within structured datasets for predictive accuracy and valuable patient health insights. Its strength with handling missing data and interpretation capabilities made it a highly promising tool in the medical field.

The addition of a Vision Transformer will be integrated within the future version of this model in an attempt to study data regarding images as much as data in tabular formats. This multi-modality can be leveraged towards more robust and strong detection methods with better overall accuracies. With that application having images of a demented brain included within its database, neurocognitive disorders are looked forward to being understood extensively at every level, providing proper detection and intervention policies well before the patient crosses some particular threshold.

References

[1] Heising, L. M., & Angelopoulos, S. (2021). “Early Diagnosis of Mild Cognitive Impairment with 2-Dimensional Convolutional Neural Network Classification of Magnetic Resonance Images”. In Proceedings of the 54th Hawaii International Conference on System Sciences.

[2] Pacheco-Lorenzo, M. R., Valladares-Rodríguez, S. M., Anido-Rifón, L. E., & Fernández-Iglesias, M. J. “Smart Conversational Agents for the Detection of Neuropsychiatric Disorders: A Systematic Review”. Department of Telematics Engineering at University of Vigo, Escola de Enxeñaría de Telecomunicación, Campus Lagoas-Marcosende, 36310 Vigo, Spain.

[3] P. Kumar, V. Subathra, Y. Swasthika and V. Vishal, "Disease Diagnosis Using Machine Learning on Electronic Health Records," 2024 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2024, pp. 1-6, doi: 10.1109/IC3IoT60841.2024.10550235.

[4] Shah, S. M. A. H., Khan, M. Q., Rizwan, A., Jan, S. U., Abdel Samee, N., & Jamjoom, M. M. “Computer-Aided Diagnosis of Alzheimer’s Disease and Neurocognitive Disorders with Multimodal Bi-Vision Transformer (BiViT”). Received: 2 August 2023 / Accepted: 14 June 2024.

[5] Sirilertmekasakul, C., Rattanawong, W., Gongvatana, A., & Srikiatkhachorn, A. “The Current State of Artificial Intelligence-Augmented Digitized Neurocognitive Screening Test”. Received: 29 December 2022 / Accepted: 20 March 2023 / Published: 30 March 2023.

[6] Ghoraani, B., Boettcher, L. N., Tolea, M. I., Galvin, J. E., Hssayeni, M. D., & Rosenfeld, A. “Detection of Mild Cognitive Impairment and Alzheimer’s Disease Using Dual-Task Gait Assessments and Machine Learning”. Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, 33431, USA / Comprehensive Center for Brain Health, Department of Neurology, University of Miami, Miami, FL, 33136, USA.

[7] Rutkowski, T. M., Abe, M. S., Koculak, M., & Otake-Matsuura, M. “Classifying Mild Cognitive Impairment from Behavioral Responses in Emotional Arousal and Valence Evaluation Task–AI Approach for Early Dementia Biomarker in Aging Societies”.

[8] Javeed, A., Dallora, A. L., Rauf, H. T., Berglund, J. S., & Anderberg, P. “Early Prediction of Dementia Using Feature Extraction Battery (FEB) and Optimized Support Vector Machine (SVM) for Classification”. Received: 15 January 2023 / Revised: 30 January 2023 / Accepted: 31 January 2023 / Published: 2 February 2023.

[9] Kamal, M. S., Northcote, A., González Crespo, R., Chowdhury, L., Dey, N., & Herrera-Viedma, E. “Alzheimer’s Patient Analysis Using Image and Gene Expression Data and Explainable-AI to Present Associated Genes”. IEEE Transactions on Instrumentation and Measurement, VOL. 70, 2021, 2513107.

[10] Tsiakiri, A., Bakirtzis, C., Plakias, S., Terzoudi, A., & Christidi, F. “Predictive Models for the Transition from Mild Neurocognitive Disorder to Major Neurocognitive Disorder: Insights from Clinical, Demographic, and Neuropsychological Data”. Received: 27 April 2024 / Revised: 27 May 2024 / Accepted: 29 May 2024 / Published: 1 June 2024.

[11] R. K. Mahendran, R. Aishwarya, R. Abinayapriya and P. Kumar, "Deep Transfer Learning Based Diagnosis of Multiple Neurodegenerative Disorders," 2024 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2024, pp. 1-4, doi: 10.1109/ESCI59607.2024.10497320.

[12] AlHarkan, K., Sultana, N., Al Mulhim, N., AlAbdulKader, A. M., Alsafwani, N., Barnawi, M., Alasqah, K., Bazuhair, A., Alhalwah, Z., Bokhamseen, D., Aljameel, S. S., Alamri, S., Alqurashi, Y., & Al Ghamdi, K.” Artificial Intelligence Approaches for Early Detection of Neurocognitive Disorders Among Older Adults”. Received: 04 October 2023 / Accepted: 29 January 2024 / Published: 16 February 2024.

[13] Lagergren, M., Fratiglioni, L., Rahm Hallberg, I., Berglund, J., Elmståhl, S., Hagberg, B., Holst, G., Rennemark, M., Sjölund, B.-M., Thorslund, M., Wiberg, I., Winblad, B., & Wimo, A. “A Longitudinal Study Integrating Population, Care and Social Services Data: The Swedish National Study on Aging and Care (SNAC)”.

[14] Luz, S. “Longitudinal Monitoring and Detection of Alzheimer’s Type Dementia from Spontaneous Speech Data”. In 2017 IEEE 30th International Symposium on Computer-Based Medical Systems.

[15] Islam, R., & Layek, M. A. “StackEnsembleMind: Enhancing Well-Being Through Accurate Identification of Human Mental States Using Stack-Based Ensemble Machine Learning”. Department of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh, 2023.

[16] A. Sivaprakasam S, S. P. S, V. SP, V. M and S. J, "A Novel Approach on Parkinson Detection with Deep Ensemble Network," 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Krishnankoil, Virudhunagar district, Tamil Nadu, India, 2024, pp. 1-5, doi: 10.1109/INCOS59338.2024.10527508.

RE-2022-424666-plag-report

ORIGINALITY REPORT

8%	6%	5%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	1%
2	fastercapital.com Internet Source	1%
3	Submitted to Karpagam Academy of Higher Education Student Paper	1%
4	pdfslide.net Internet Source	1%
5	Submitted to University of Adelaide Student Paper	<1%
6	Submitted to University of Pittsburgh Student Paper	<1%
7	www.ijrte.org Internet Source	<1%
8	"Bioinformatics Research and Applications", Springer Science and Business Media LLC, 2021 Publication	<1%

CO-PO Mapping

PROJECT WORK COURSE OUTCOME (COs):

CO1: On completion it will prove as a major breakthrough in digital transformation of college management leveraging end-to-end technologies.

CO2: It will ease out the management overhaul and boost better transparency and robustness to the entire setup.

CO3: Given the huge amount of data available in the educational sector, especially in the colleges, technologies like Machine Learning and AI can be used to increment student performance and job-market ready.

CO4: It helps in keeping the entire system snappy and ensures all endpoints are taken care of, reducing the overall waiting periods in the traditional working.

CO5: Students will be able to publish or release the project to society.

PROGRAM OUTCOMES (POs)

PO1: Engineering Knowledge: Apply the knowledge of engineering fundamentals, mathematics, science and technology and an engineering specialization to the solution of complex engineering problems.

PO2: Problem analysis: Ability to apply deep learning methodologies to solve computational tasks, model real world problems using appropriate datasets and suitable deep learning models. To understand standard practices and strategies in software project development using open-ended programming environments to deliver a quality product.

PO3: Design/development of solutions: Design solution for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety.

PO4: Conduct investigations of complex problems: Use research - based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis the information to provide valid conclusions.

PO5: Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6: The Engineer and society: Apply reasoning informed by the contextual knowledge to assess social, health and safety issues and the consequent responsibilities relevant to the professional engineering practice.

PO7: Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental context, and demonstrate the knowledge of, and need for sustainable development.

PO8: Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practices.

PO9: Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10: Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11: Project management and finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one’s own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO12: Life-long learning: Recognize the need for and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAM SPECIFIC OUTCOMES (PSOs):

PSO1: Foundation Skills: Ability to understand, analyze and develop computer programs in the areas related to algorithms, system software, web design, deep learning and cloud computing for efficient design of computer-based systems of varying complexity. Familiarity and practical competence with a broad range of programming languages and open-source platforms.

PSO2: Problem-solving Skills: Ability to apply mathematical methodologies to solve computational tasks, model real world problems using appropriate data structure and suitable algorithms. To understand standard practices and strategies in software project development using open-ended programming environments to deliver a quality product.

PSO3: Successful Progression: Ability to apply knowledge in various domains to identify research gaps and to provide solutions to new ideas, inculcate passion towards higher studies, creating innovative career paths to be an entrepreneur and evolving as an ethically responsible computer science professional.