ISSN No:-2456-2165

Early Cognitive Changes and Predictive Models: A Multilayer Perceptron with Neuropsychological Test Data

Ibrahim Almubark
Department of Information Technology
College of Computer, Qassim University
Buraydah, Saudi Arabia

Abstract:- The capacity to predict the seemingly ambiguous transition from mild cognitive impairment (MCI) to progressive cognitive decline is a critical concern in cognitive research. Advancement in computational systems has contributed to more robust potential to apply innovations in this sector. This study uses a multilayer perceptron (MLP) neural network approach to investigate and compare the utility of various neuropsychological tests to predict a 3-year progression from MCI. The MLP neural network is developed using the open database from the Alzheimer's Disease Neuroimaging Initiative (ADNI). The data were based on a sample of 246 subjects with MCI whose diagnostic follow-up was available for at least the full 3-year period after the initial baseline assessment during the initial project period, i.e., ADNI-1. Classification results and analysis demonstrated that the combined features from all three neuropsychological tests outperformed a single test and the pairwise tests with an accuracy of 89.43%, a sensitivity of 89.19%, a specificity of 89.63%, and the area under the receiver operating characteristic curve (AUC) of 0.934.

Keywords:- *Mild Cognitive Impairment; Artificial Neural Network; Neuropsychological Testing.*

I. INTRODUCTION

Dementia is the umbrella term used to describe a decline in cognitive ability that has a significant impact upon an individual's functioning related to the undertaking of the tasks of everyday life. Globally, approximately 47 million individuals are currently living with dementia [1], while 24 million are living with Alzheimer's disease (AD) [2]. As the mean age of the global population increases, so do rates of AD, and it is currently predicted that prevalence will increase four-fold by 2050 [2]. Caring for those with AD is resource intensive, with $305 billion spent on the care of those with disease in the U.S. in 2020 [3]. Before the development of diagnosable AD, patients experience mild cognitive impairment (MCI) that can have an impact upon their everyday functioning. Identification of MCI is crucial because it represents the earliest clinically detectable stage of a potential progression toward various dementias, including AD [4]. However, not all MCI patients transition to AD, thus it is essential that we develop an effective method that can accurately distinguish between MCI subjects who will progress to AD and those who will remain stable/show improvement.

The availability of large biomedical datasets has increased in recent years, accompanied by advances in the field of artificial intelligent (AI) technologies. In combination, these developments have improved and increased our ability to diagnose AD [5], [6]. Current research has tended to focus upon the utility of different combinations of biomarkers to predict the conversion of AD in MCI patients. Such biomarkers include brain imaging data, cerebrospinal fluid (CSF) specimens, genotyping, and neuropsychological tests [7], [8]. Neuropsychological tests also appear to be feasible and effective for disease prognosis, while also being relatively inexpensive to administer. The undertaking of neuropsychological tests includes the investigation of cognitive processes and functioning involved in corresponding tasks such as thinking, planning, walking, remembering, talking, seeing, and feeling [3]. A decline in the cognitive functioning, required to undertake these tasks, has the ability to drastically reduce an individual's quality of life and is linked to higher rates of morbidity and mortality [9], [10]. Therefore, the presence of multiple cognitive deficits suggests that the efficacy of using a combination of neuropsychological tests from various domains to characterize the patterns developed due to cognitive impairments and enable improved clinical diagnosis [11]–[13].

Based on the subsequent diagnosis status at follow-up assessments, MCI patients can be divided into two subgroups: (1) subjects diagnosed with MCI who remain stable (defined as stable MCI (sMCI)) and (2) subjects who progress to AD (defined as progression MCI (pMCI)). The current study sought to develop fully connected multilayer perceptron (MLP) neural networks to predict whether an MCI patient would progress to AD during a 3-year follow-up period (i.e., sMCI vs. pMCI) based on the itemized scores from three neuropsychological tests contained within the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Initial results were compared with the state-of-the-art models used within more complex methodologies.

The paper is organized as follows. Section II presents the data and the methodology of the MLP. Section III presents the results of the modeling. Section IV evaluates and discusses

the results from our research. The final section derives conclusions from the study and proposes directions for future work.

II. MATERIALS AND METHODS

Analyses were implemented utilizing Python and several associated libraries such as *Scikit-learn*, *Pandas*, *Numpy*, *TensorFlow*, and *Keras* [14], [15]. The following steps were undertaken to construct the classifiers and optimize their performance. A summary of all the steps involved in building a classification model is represented in Fig. 1.

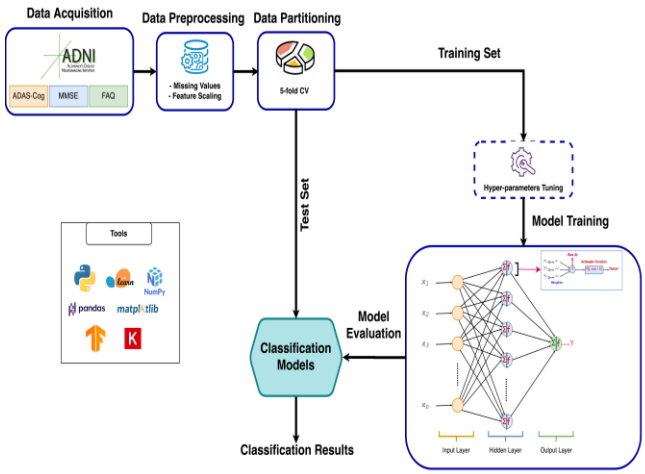

Fig. 1. Steps involved in building a classification model.

A. ADNI Database

The ADNI database (http://adni.loni.usc.edu/) is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of AD progression. Briefly, the database includes subjects recruited from over 50 sites across the U.S. and Canada and draws on a broad range of academic institutions and private corporations. Led by Principal Investigator Michael W. Weiner, MD, the project began in 2003 and has been extended to different phases. The first phase of ADNI (ADNI-1) was completed in 2010, followed by ADNI-GO, ADNI-2, and ADNI-3. These four protocols have recruited over 1,900 adults between the ages of 55 to 90 years. The sample includes those that are cognitively normal (CN), living with MCI, and individuals with AD. The follow-up duration of each group is described in the protocols for ADNI-1, ADNI-GO, ADNI-2, and ADNI-3 (www.adni-info.org).

B. Subjects

In this study, we used the baseline visit data from 391 participants with MCI recruited during the initial project period (ADNI-1). Patients who were diagnosed with MCI at all visits during the 3-year follow-up period were included in the sMCI group. Patients whose diagnosis changed to AD during the follow-up period were included in the pMCI group. In total, 135 participants were in the sMCI group and 111 in the pMCI group; 145 were lost to follow-up.

Participants were aged between 55-90, in good overall health and had no evidence of cerebrovascular disease. Further inclusion criteria included having at least six years of education or work history and fluency in either English or Spanish. Every subject, along with their partners, completed the informed consent process. The study protocols underwent review and approval by the Institutional Review Board at each ADNI data collection site. Table I displays the characteristics of the sMCI and pMCI subjects involved in this study. The mean test score was computed by averaging the scores from all the question in one test (see section *C*)

TABLE I. BASELINE VISIT CHARACTERISTICS OF SUBJECTS RECRUITED DURING ADNI-1.

Characteristic	sMCI (n=135)	pMCI (n=111)	*p*-value
Age, year	74.28 ± 1.27	74.61 ± 1.26	0.72
Education, years	15.66 ± 0.50	15.77 ± 0.56	0.76
Sex, male/female	94/41	72/39	0.43
ADAS-Cog score	15.48 ± 1.00	20.81 ± 1.00	3.70×10^{-12}
MMSE score	27.64 ± 0.29	26.67 ± 0.31	1.09×10^{-5}
FAQ score	1.90 ± 0.51	5.86 ± 0.94	6.54×10^{-13}

Values are shown as mean and the 95% confidence interval or gender ratios. The *p*-values for differences between sMCI and pMCI are based on one-way ANOVA test. ADAS-Cog = Alzheimer's Disease Assessment Scale-Cognitive subscale; MMSE = Mini-Mental State Examination; FAQ = Functional Activities Questionnaire.

Before creating a neural network model, a one-way ANOVA test was performed to determine if the differences in the mean between the two groups (sMCI and pMCI) are statistically significant. From Table I, the mean differences between test scores are all statistically significant, while the mean differences in the demographic variables are not statistically significant (*p-value* > 0.05).

C. Neuropsychological Data

Three neuropsychological assessments from the ADNI-1 baseline visit, containing Alzheimer's Disease Assessment Scale-Cognitive subscale (ADAS-Cog) [16], Mini-Mental State Examination (MMSE) [17], and Functional Activities Questionnaire (FAQ) [18], were used. The scores of individual questions were assessed (Table II), there were 13, 30, and 10 itemized scores in ADAS-Cog, MMSE, and FAQ, respectively. The Q13 in ADAS-Cog was not used by default. Each score was derived as a summary result for a set of answers given for that test and was then included as a feature in the machine learning task.

D. Designing Experiments for Classification

The MLP neural networks were constructed utilizing: (1) the original set of features from each individual neuropsychological test (described above); (2) the original set of 43 features from ADAS-Cog and MMSE; (3) the original set of 23 features from ADAS-Cog and FAQ; (4) the original set of 40 features from MMSE and FAQ; and (5) the set of the 53 features from the three neuropsychological tests combined. We trained seven MLP models to perform classification between sMCI vs. pMCI for both the individual, pairwise, and combined tests.

TABLE II. NEUROPSYCHOLOGICAL ASSESSMENTS EMPLOYED IN THIS STUDY.

Neuropsychological Tests	
ADAS-Cog	Registration (3)
Q1. Word recall	Attention and calculation (5)
Q2. Word recognition	Recall (3)
Q3. Object naming	Language (8)
Q4. Recall test instructions	Visual construction (1)
Q5. Orientation	**FAQ**
Q6. Commands	Q1. Manage finances
Q7. Clarity of language	Q2. Complete forms
Q8. Comprehension	Q3. Shop
Q9. Word finding	Q4. Perform games of skill or hobbies
Q10. Ideational praxis	Q5. Prepare hot beverages
Q11. Constructional praxis	Q6. Prepare a balanced meal
Q12. Delayed word recall	Q7. Follow current events
Q14. Number cancellation	Q8. Attend to TV, books, or magazines
MMSE	Q9. Remember appointments
Orientation (10)	Q10. Travel out of the neighborhood

E. Data Preprocessing

The baseline dataset in ADNI-1 comprises of 400 subjects diagnosed with MCI. The count of MCI subjects with missing values was notably low (9 out of 400) and we, therefore, chose to remove those subjects without using a replacement technique. Therefore, the dataset (n=391) used in our study is free of any missing values. In total, 135 participants were in the sMCI group and 111 in the pMCI group; 145 were lost to follow-up. In statistics, a common practice involves discarding cases with missing values if they represent less than 5% of the total samples, provided the overall sample size is substantial enough.

Feature normalization was executed through standard scaling, achieving a mean of Zero and a standard deviation of one. Feature normalization accelerates and stabilizes the optimization process [19]. Machine learning algorithms typically possess one or more hyper-parameters that significantly influence the model's performance. Hyper-parameters can be adjusted until the optimal model is found, a process known as hyper-parameters optimization/tuning. It is important to find and select the best hyper-parameters because they determine how learning of the algorithm is performed and controlled. In this study, prior to fitting the training data to each model, a grid search was performed to acquire the optimal set of hyper-parameters for the MLP networks. The grid search operates by systematically exploring a defined subset of hyper-parameters to identify the most optimal combination for a given network [20]. The best hyper-parameters are chosen based on the area under the receiver operating characteristic curve (AUC) calculated for each combination of parameters.

The evaluation of the predictive model was performed using cross-validation (CV) with stratified K-Fold. Data was divided into five subsets with consistent ratios between classes in each fold. In each fold, 80% of the data was utilized for training while the remaining 20% was allocated for testing. The sensitivity, specificity, accuracy, and AUC were calculated from the 5-fold CV.

F. Multilayer Perceptron (MLP)

The MLP is a feed-forward artificial neural network employing the back-propagation algorithm for weight updates [21]. The term "feed-forward" relates to the information from the input nodes going through the hidden nodes and moving forward to the output nodes. Each node is a neuron (processing unit) with a non-linear activation function. MLP has several advantages to perceptrons, in that it is able to identify linearly indivisible data, a task not possible using perceptrons [22].

The architecture of MLP is comprised of an input layer, one or more hidden layers, and an output layer as shown in Fig. 1. Within the current study, we adopted a three-layer architecture MLP neural network to perform the classification tasks. The dimension of the input vector, in the input layer, depends on the number of input features. We adopted a single hidden layer, and then performed hyper-parameter tuning using a different number of hidden neurons to define the optimal number of neurons in the hidden layer. The output layer has only one neuron since it only has two classes (sMCI and pMCI) within it.

The rectified linear unit (ReLU) was selected as the activation function for the input and hidden layers [23]. As ReLU is a linear function and it takes a value of zero for all negative values, it can be defined as [24]:

$$y = max(0, x) \tag{1}$$

We also used the sigmoid function as the activation function for the output layer to obtain output between 0 and 1 for prediction of probabilities. The sigmoid function follows an s-shaped curve and can be defined as [25]:

$$S(x) = \frac{e^x}{e^x + 1} \tag{2}$$

The optimization of the networks was performed using the adaptive moment estimation (Adam) algorithm. This is an adaptive learning rate optimization algorithm designed for training neural networks and has demonstrated high efficiency [26]. The maximum number of epochs was set to 250 to ensure the training set loss function converges within a tolerance of 10^{-4}. Both the *L1* and *L2* regularization were added to each layer to reduce the possibility of overfitting. Early stopping and learning rate shrinkage (with a minimum learning rate of 5×10^{-4}) was performed to monitor the validation loss function. Binary cross entropy was used as the loss function. The weight ratio between the two classes in the loss function was treated as one of the tunable hyper-parameters. After getting the probability for each sample from the trained network, and in order to get the highest AUC, we treated the threshold for classifying each sample to sMCI or pMCI as a hyper-parameter. Practically, one can tune the thresholds and the class weights to improve the score of other metrics, such as overall accuracy, or True Positive Rate (TPR) such that Positive Predict Value (PPV) meets an agreed upon requirement. Our work shows the tuned results for the AUC as an example (discussed in Section III).

G. Performance Assessment

For classifier assessment, sensitivity, specificity, and accuracy were calculated for each model (equations 3, 4, and 5 below, respectively). Sensitivity, referred to as recall, assesses the proportion of true positive subjects identified by the test among all subjects identified as positive. The values of sensitivity are expressed within the 0 to 1 range, with higher values indicative of a higher value of true positives and a lower value of false negatives. Specificity gauges the true negative rate by measuring the proportion of actual negative subjects among the total number of subjects testing negative. As with sensitivity, values closer to one are preferable. Accuracy is the ratio of correctly classified subjects to entire subjects and represents one of the most important metrics in machine learning. Formulas are given below, where: TP = True Positives; TN = True Negatives; FN = False Negatives; FP = False Positives.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (4)$$

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

TABLE III. MULTILAYER PERCEPTRON (MLP) CLASSIFICATION PERFORMANCE USING DATA FROM EACH INDIVIDUAL TEST, THE ORIGINAL SET OF FEATURES FROM ADAS-COG AND MMSE, THE ORIGINAL SET OF FEATURES FROM ADAS-COG AND FAQ, THE ORIGINAL SET OF FEATURES FROM MMSE AND FAQ, AND THE COMBINED-TEST TO CLASSIFY SMCI VS. PMCI. THE SENSITIVITY, SPECIFICITY, ACCURACY, AND AUC WERE CALCULATED FROM 5-FOLD CV USING THE DEFAULT SETTING FOR CLASS WEIGHT (1:1) AND PROBABILITY THRESHOLD (0.5). THE PERFORMANCE WITH THE OPTIMAL HYPER-PARAMETER TUNING IS SHOWN IN BOLD FONTS (OPTIMAL VALUES FOR CLASS WEIGHT AND PROBABILITY THRESHOLD).

Case	Dataset	Probability Threshold	Class Weight	Sensitivity %	Specificity%	Accuracy %	AUC
sMCI vs. pMCI	ADAS-Cog (13)	0.5	1:1	63.39	79.56	72.29 ± 6.95	0.795
		0.43	**1:0.83**	**81.25**	**70.07**	**75.10 ± 7.30**	**0.829**
	MMSE (30)	0.5	1:1	52.21	78.99	66.93 ± 6.82	0.677
		0.56	**1:1**	**44.25**	**92.75**	**70.92 ± 8.54**	**0.790**
	FAQ (10)	0.5	1:1	65.18	78.68	72.58 ± 2.22	0.793
		0.43	**1:0.66**	**66.96**	**81.62**	**75.00 ± 3.37**	**0.805**
	ADAS-Cog + MMSE (43)	0.5	1:1	75.00	87.83	77.11 ± 8.23	0.848
		0.3	**1:1.83**	**98.21**	**52.55**	**73.09 ± 5.97**	**0.898**
	ADAS-Cog + FAQ (23)	0.5	1:1	63.06	88.15	76.83 ± 5.46	0.828
		0.46	**1:1.33**	**81.98**	**79.26**	**80.49 ± 5.81**	**0.891**
	MMSE + FAQ (40)	0.5	1:1	68.75	86.76	78.63 ± 7.41	0.850
		0.6	**1:1.33**	**67.86**	**89.71**	**79.84 ± 6.98**	**0.885**
	Combined-Test (53)	0.5	1:1	79.28	85.19	82.52 ± 4.04	0.900
		0.46	**1:1.16**	**89.19**	**89.63**	**89.43 ±10.28**	**0.934**

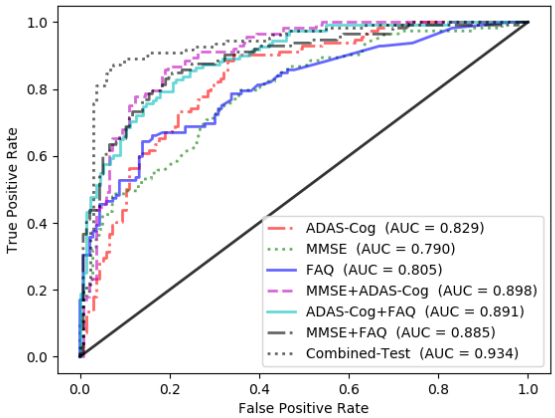

Fig. 2. The ROC curves using data from each individual test, the original set of features from ADAS-Cog and MMSE, the original set of features from ADAS-Cog and FAQ, the original set of features from MMSE and FAQ, and the combined-test with multilayer perceptron (MLP). AUC score shown in the legend box.

Based on the predicted probabilities for each of our participants, we assessed and plotted the Receiver Operating Characteristic (ROC) curve and calculated the Area Under the Curve (AUC). ROC/AUC is the most widely used metrics for evaluating a binary classifier, and is considered to be a reliable metric when a dataset has unbalanced counts for different target classes. For ROC curves, plotting involves an x-axis measuring the False Positive Rates (FPR) and a y-axis measuring the True Positive Rates (TPR). As the AUC moves closer to one, the performance of a predictive model in relation to reliability is assessed to be improving – with one being indicative of a perfect model performance.

III. RESULTS

Table III summarizes the MLP classification performance obtained from using each individual test (13 from ADAS-Cog, 30 from MMSE, and 10 from FAQ), the original set of features from ADAS-Cog and MMSE (43 features), the original set of features from ADAS-Cog and FAQ (23 features), the original set of features from MMSE and FAQ (40 features), and the combined-test of 53 features from the three neuropsychological tests to predict whether an MCI patient, who classified as MCI at the baseline visit, would progress to AD during follow-up. The average classification performance using the standard class weight (1:1) and probability threshold (0.5) is presented, alongside the model's performance with an optimal setting (highlighted in bold). Each optimal setting was determined by tuning class weight and threshold binarizing within the neural network to increase the performance via grid search tuning. As displayed in Table III, the model utilizing combined features, from the three tests, demonstrated superior performance compared to models using single test and the pairwise tests features. In particular, the combined-test model, with optimal parameters, obtained an AUC value of 0.934, which indicates high accuracy (few false negative and false positive cases). The models using pairwise tests outperformed the models using single test features, which also stresses that the combination of neuropsychological tests from various domains could improve classification performance.

Fig. 2 plots the ROC curves using MLP for each individual test, pairwise tests, and the combined-test. The AUC values suggest that the best model is the combined-test model, followed by the ADAS-Cog + MMSE, ADAS-Cog + FAQ, MMSE + FAQ, ADAS-Cog, FAQ, and MMSE.

IV. DISCUSSION

Utilizing data from three neuropsychological tests, our current study constructed several, fully connected MLP neural networks to classify whether an MCI patient was likely to progress from MCI to AD during a 3-year follow-up period.

Despite there being other neuropsychological tests, the three included in the current study were selected for the following reasons. Firstly, since a prominent feature of AD is memory impairment, ADAS-Cog and MMSE were selected as appropriate tests for this study. Secondly, ADAS-Cog and MMSE are tests of global cognitive function and, therefore, cover several domains outside of memory [27]. Thirdly, though designed for subjects with AD, the ADAS-Cog test demonstrates efficacy when used as a measure for trials of interventions in subjects with MCI [28] and is widely used as a cognitive scale in clinical trials [29]. Finally, as functional changes are noted earlier in the dementia process [30], [31], data on function from the FAQ test were also included. The FAQ has been found to be consistently accurate and demonstrates good sensitivity and specificity [18], [32]. The FAQ also has the ability to distinguish between different cognitive groups and, in particular, distinguish MCI from mild AD [33], [34].

Discriminating and classifying sMCI versus pMCI is a particularly difficult and challenging task because their shared key features often appear to overlap and have a number of key similarities [35], [36]. Our classification results outperformed earlier research in this area, including the work of Grassi et al. [37] by 11% accuracy. Within their study they utilized clinical and neuropsychological test scores, cardiovascular risk indexes, and a visual rating scale for brain atrophy. The increased efficacy of our model to distinguish between these two groups supports its potential use.

Our approach used only neuropsychological tests and appears to have outperformed more complex study methodologies – including those which used biomarkers such as brain imaging and CSF. The use of such biomarkers can be prohibitively expensive and/or invasive and, because of this, is often not offered in the clinical setting.

In comparison to the previous research conducted by Pang et al. [38] and Massetti et al. [39], our approach focused solely on neuropsychological tests and achieved a notable accuracy of up to 89% in predicting outcomes. Pang et al. [38] utilized a variety of clinical variables from the National Alzheimer's Coordinating Center (NACC) dataset to predict transitions from normal cognition to MCI and from MCI to AD using machine learning classifiers such as Support Vector Machines (SVM), Logistic Regression (LR), and Random Forest (RF). Their findings revealed that for the prediction of progression to AD within 3 years, neuropsychological tests, memory, community affairs, and judgement subitems played a significant role, whereas for the 2-year follow-up, biomarkers were more influential. The highest accuracies, ranging from 80% to 85%, were achieved using the RF model for both 2-year and 3-year prediction periods.

Massetti et al. [39] on the other hand, utilized data from the ADNI and Alzheimer's Disease Metabolomics Consortium (ADMC) datasets to predict MCI-to-AD transitions. They employed the RF algorithm on a dataset of 587 MCI subjects, considering neuropsychological test scores, AD related cerebrospinal fluid (CSF) biomarkers, peripheral biomarkers, and structural MRI data as variables. Their results indicated an accuracy of 86% in predicting the MCI-to-AD transition, with neuropsychological test scores, MRI data, and CSF biomarkers identified as the most crucial features.

Additionally, Diogo et al. [40] focused on early detection of AD using MRI scans from ADNI and Outcome and Assessment Information Set (OASIS) databases. Their study included 570 subjects from ADNI and 531 subjects from the OASIS dataset, aiming to classify healthy controls (HC), MCI, and AD using an ensemble machine learning model. Morphometric and graph theory features were extracted from MRI scans for analysis. Their classification tasks employed various machine learning techniques such as linear SVM, decision tree, random forest, extremely randomized tree, linear discriminant analysis, logistic regression, and logistic regression with stochastic gradient descent learning. The study achieved a balanced accuracy of 90.6% for HC vs. AD classification and 62.1% for HC vs. MCI vs. AD classification.

In the context of these comprehensive studies, our approach stands out due to its singular reliance on neuropsychological tests, which suggests that a focused approach on these tests might offer a more streamlined and effective predictive model compared to the integration of broader variable ranges, as demonstrated by previous research. It is important to consider the distinct datasets, methodologies, and prediction timeframes employed in these studies, which may contribute to the variability in accuracy and feature importance observed across the research findings. Further exploration and validation of our approach could provide valuable insights into its applicability and significance in predicting cognitive transitions.

Our classification model achieved a more optimal performance-related result compared to earlier research that often utilized more complex methodologies [37]–[40]. Our study demonstrates the potential of incorporating neuropsychological tests into regular physical examinations for seniors. At the same time, early detection of AD becomes feasible when these tests are used in combination with deep learning. This study also validates that utilizing a combination of a multiple neuropsychological tests and assessments improve the accuracy of clinical diagnosis in AD. This approach is currently used in practice, since a physician can miss a diagnosis due to the utilization of a single test, or one that is not sensitive to differences between MCI and AD. Moreover, early detection of MCI is critical to putting in place treatment regimens able to slow cognitive decline and reduce its impact upon quality of life.

V. CONCLUSIONS AND FUTURE WORK

The AD is a progressive disorder that incorporates a range of symptoms, and its insidious nature means that it develops over time, often beginning with very mild symptoms that can easily be mistaken or ignored. Consequently, there is a pressing need to develop accurate and reliable early prediction models capable of detecting potential changes from mild cognitive impairment (MCI) to AD in a timely manner. In this pursuit, the current study employed MLP neural networks, which demonstrated good results by achieving an accuracy rate of close to 90%. This result highlights the potential of deep learning-enabled classifiers to effectively discriminate between individuals at varying risk levels for AD progression. Such advancements in artificial intelligence-driven methodologies can significantly aid in identifying individuals who are in the early stages of AD, thus allowing for more timely interventions and personalized treatment strategies.

The implications of these findings are substantial. Firstly, early detection of AD can lead to early interventions that may help slow down the disease's progression and improve the quality of life for affected individuals. Moreover, the accurate identification of individuals at higher risk of AD could enable more targeted research and clinical trials, ultimately facilitating the development of novel therapeutics to combat the disease.

While the current study's results are promising, further research is warranted to validate and refine the findings, especially in diverse clinical settings. Future investigations could explore larger and more diverse cohorts to ensure the generalizability of the predictive model. Additionally, longitudinal studies can shed light on the model's performance over an extended period, helping to ascertain its long-term accuracy and reliability.

In conclusion, the development of an accurate early prediction model for AD is of paramount importance in tackling this devastating disease. The success of the MLP neural network in discriminating between groups at different stages of AD risk represents a significant step forward in this endeavor. With continued research and refinement, these innovative AI-driven approaches hold immense potential in revolutionizing how we diagnose and manage AD, ultimately making a positive impact on the lives of millions of individuals and their families worldwide.

ACKNOWLEDGMENT

The author would like to thank Qassim University and the Deanship of Scientific Research for their support and funding the publication.

REFERENCES

[1]. A. Ponjoan *et al.*, "Epidemiology of dementia: prevalence and incidence estimates using validated electronic health records from primary care," *Clinical epidemiology*, vol. 11, p. 217, 2019.
[2]. A. Kumar and J. W. Tsao, "Alzheimer disease," 2019.
[3]. Alzheimer's Association, "2020 Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 16, no. 3, pp. 391–460, 2020.
[4]. K. Ritter *et al.*, "Multimodal prediction of conversion to Alzheimer's disease based on incomplete biomarkers," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 1, no. 2, pp. 206–215, 2015.
[5]. A. L. Dallora, S. Eivazzadeh, E. Mendes, J. Berglund, and P. Anderberg, "Machine learning and microsimulation techniques on the prognosis of dementia: A systematic literature review," *PloS one*, vol. 12, no. 6, p. e0179804, 2017.
[6]. T. Jo, K. Nho, and A. J. Saykin, "Deep Learning in Alzheimer's disease: Diagnostic Classification and Prognostic Prediction using Neuroimaging Data," *arXiv preprint arXiv:1905.00931*, 2019.
[7]. J. Dukart, F. Sambataro, and A. Bertolino, "Accurate prediction of conversion to Alzheimer's disease using imaging, genetic, and neuropsychological biomarkers," *Journal of Alzheimer's Disease*, vol. 49, no. 4, pp. 1143–1159, 2016.
[8]. H.-I. Suk, S.-W. Lee, D. Shen, and A. D. N. Initiative, "Latent feature representation with stacked auto-encoder for AD/MCI diagnosis," *Brain Structure and Function*, vol. 220, no. 2, pp. 841–859, 2015.
[9]. K. M. Langa *et al.*, "Trends in the prevalence and mortality of cognitive impairment in the United States: is there evidence of a compression of cognitive

morbidity?," *Alzheimer's & Dementia*, vol. 4, no. 2, pp. 134–144, 2008.
[10]. S. van Roessel, C. J. Keijsers, and M. D. Romijn, "Dementia as a predictor of morbidity and mortality in patients with delirium," *Maturitas*, vol. 125, pp. 63–69, 2019.
[11]. P. D. Harvey, "Clinical applications of neuropsychological assessment," *Dialogues in clinical neuroscience*, vol. 14, no. 1, p. 91, 2012.
[12]. E. Storey and G. J. Kinsella, "Principles of neuropsychometric assessment," in *Neurology and Clinical Neuroscience*, Mosby, 2007, pp. 22–30.
[13]. S. D. Yeatts, Y. Y. Palesch, and N. Temkin, "Biostatistical issues in TBI clinical trials," in *Handbook of Neuroemergency Clinical Trials*, Elsevier, 2018, pp. 167–185.
[14]. F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.
[15]. F. Chollet, "Keras: Deep learning library for theano and tensorflow," *URL: https://keras. io/k*, vol. 7, no. 8, p. T1, 2015.
[16]. W. Rosen, R. Mohs, and K. Davis, "Alzheimer's Disease Assessment Scale—Cognitive and Non-Cognitive Sections (ADAS-Cog, ADAS Non-Cog)," *Journal of Psychiatry*, vol. 141, pp. 1356–1364, 1984.
[17]. M. F. Folstein, S. E. Folstein, and P. R. McHugh, "'Mini-mental state': a practical method for grading the cognitive state of patients for the clinician," *Journal of psychiatric research*, vol. 12, no. 3, pp. 189–198, 1975.
[18]. R. I. Pfeffer, T. T. Kurosaki, C. H. Harrah Jr, J. M. Chance, and S. Filos, "Measurement of functional activities in older adults in the community," *Journal of gerontology*, vol. 37, no. 3, pp. 323–329, 1982.
[19]. S. Raschka and V. Mirjalili, *Python machine learning*. Packt Publishing Ltd, 2017.
[20]. J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, no. Feb, pp. 281–305, 2012.
[21]. S. Marsland, *Machine learning: an algorithmic perspective*, 2nd ed. CRC press, 2015.
[22]. G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 5, no. 4, pp. 455–455, 1992.
[23]. B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *arXiv preprint arXiv:1505.00853*, 2015.
[24]. B. Hanin, "Universal function approximation by deep neural nets with bounded width and relu activations," *Mathematics*, vol. 7, no. 10, p. 992, 2019.
[25]. J. Han and C. Moraga, "The influence of the sigmoid function parameters on the speed of backpropagation learning," in *International Workshop on Artificial Neural Networks*, 1995, pp. 195–201.
[26]. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.
[27]. R. Casanova *et al.*, "Alzheimer's disease risk assessment using large-scale machine learning methods," *PLoS One*, vol. 8, no. 11, p. e77949, 2013.
[28]. J. Skinner *et al.*, "The Alzheimer's disease assessment scale-cognitive-plus (ADAS-Cog-Plus): an expansion of the ADAS-Cog to improve responsiveness in MCI," *Brain imaging and behavior*, vol. 6, no. 4, pp. 489–501, 2012.
[29]. D. J. Connor and M. N. Sabbagh, "Administration and scoring variance on the ADAS-Cog," *Journal of Alzheimer's Disease*, vol. 15, no. 3, pp. 461–464, 2008.
[30]. P. J. Brown, D. P. Devanand, X. Liu, and E. Caccappolo, "Functional impairment in elderly patients with mild cognitive impairment and mild Alzheimer disease," *Archives of general psychiatry*, vol. 68, no. 6, pp. 617–626, 2011.
[31]. S. E. John, A. S. Gurnani, C. Bussell, J. L. Saurman, J. W. Griffin, and B. E. Gavett, "The effectiveness and unique contribution of neuropsychological tests and the δ latent phenotype in the differential diagnosis of dementia in the uniform data set.," *Neuropsychology*, vol. 30, no. 8, p. 946, 2016.
[32]. M. Malek-Ahmadi *et al.*, "Sensitivity to change and prediction of global change for the Alzheimer's Questionnaire," *Alzheimer's research & therapy*, vol. 7, no. 1, p. 1, 2015.
[33]. E. Teng, B. W. Becker, E. Woo, D. S. Knopman, J. L. Cummings, and P. H. Lu, "Utility of the Functional Activities Questionnaire for distinguishing mild cognitive impairment from very mild Alzheimer's disease," *Alzheimer disease and associated disorders*, vol. 24, no. 4, p. 348, 2010.
[34]. G. A Marshall *et al.*, "Functional activities questionnaire items that best discriminate and predict progression from clinically normal to mild cognitive impairment," *Current Alzheimer Research*, vol. 12, no. 5, pp. 493–502, 2015.
[35]. D. S. Knopman and R. C. Petersen, "Mild cognitive impairment and mild dementia: a clinical perspective," in *Mayo Clinic Proceedings*, 2014, vol. 89, no. 10, pp. 1452–1459.
[36]. R. O. Roberts *et al.*, "Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal," *Neurology*, vol. 82, no. 4, pp. 317–325, 2014.
[37]. M. Grassi, D. A. Loewenstein, D. Caldirola, K. Schruers, R. Duara, and G. Perna, "A clinically-translatable machine learning algorithm for the prediction of Alzheimer's disease conversion: further evidence of its accuracy via a transfer learning approach," *International psychogeriatrics*, vol. 31, no. 7, pp. 937–945, 2019.
[38]. Y. Pang *et al.*, "Predicting Progression from Normal to MCI and from MCI to AD Using Clinical Variables in the National Alzheimer's Coordinating Center Uniform Data Set Version 3: Application of Machine Learning Models and a Probability Calculator," *The journal of prevention of Alzheimer's disease*, vol. 10, no. 2, pp. 301-313, 2023.
[39]. N. Massetti *et al.*, "A Machine Learning-Based Holistic Approach to Predict the Clinical Course of Patients within the Alzheimer's Disease Spectrum 1," *Journal of Alzheimer's Disease*, vol. 85, no. 4, pp. 1639-1655, 2022.

[40]. V. S. Diogo, H. A. Ferreira, D. Prata, and A. D. N. Initiative, “Early diagnosis of Alzheimer’s disease using machine learning: a multi-diagnostic, generalizable approach,” *Alzheimer's Research & Therapy*, vol. 14, no. 1, pp. 107, 2022.