ISSN 2347 - 3983
Volume 8. No. 7, July 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter 13872020.pdf
https://doi.org/10.30534/ijeter/2020/13872020

Automated Prognostic Modelling of Alzheimer's Disease Prediction based on Machine Learning over Brain Networks

P. S. Ramesh[1], S. Arivalagan[2], P. Sudhakar[3]

Research Scholar, Annamalai University, Annamalai Nagar, Chidambaram, Tamilnadu, India. ramesh.ps@yahoo.com
[2]Assistant Professor, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Chidambaram, Tamilnadu, India. arivucseau@gmail.com
[3]Associate Professor, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Chidambaram, Tamilnadu, India. kar.sudha@gmail.com

ABSTRACT

Medical Institutions rely on implementing E-Systems for tracking, monitoring and simplifying the process of diagnosing the life threatening diseases. Many advanced electronic health monitoring systems have become sophisticated with immense development in computing and communication technologies. Machine Learning is another technique that has a superior command over image processing, diagnosis and disease predictions. This research article emphasizes on the technique to provide an intelligent and novel technique to predict the progress of Alzheimer's disease through a decision support system. The original raw images from the MRI devices are converted into a functional matrix which illustrates the 90 credible brain regions and thus their activities. Once the images are tracked down to their images and operational centres, the association between various regions is represented as a correlation matrix of the functional areas. This correlation process is computed for every pair of brain regions from the overall images. An automated encoder network is built to identify and classify the regions affected in close relation to the Alzheimer's disease. The experimental set up evaluates the proposed mechanism against the conventional models of prediction and the results reveal that the proposed scheme is faster and reliable than the conventional strategies. The correlations between different regions of the brain achieved with an improved accuracy to detect the Alzheimer's disease. This proposal has shown significant increase of 25% betterments when compared to Support Vector Machine (SVM). This benefit has improvised the standards of electronic support systems in medical industries and prediction of Alzheimer's Disease at early stages facilitates the diagnosis process pre-emptively. The onset of Alzheimer's Disease should be limited to primary stages provided the method is applicable in real life applications.

Key words : Alzheimer's Disease, Prediction system, deep learning, machine learning, intelligent systems.

1. INTRODUCTION

Advancements in the medical industry has attained great heights as the practice of maintaining Electronic Health Records (HER) and automated applications. This technology is supported by various combinations of electronic processes and communication technologies. The applications vary from Electronic Medical Records (EMR) or EHRs [1], clinical decision support systems, virtual doctoring systems, telemedicine [2] and Healthcare Information Systems [3]. This approach has facilitated the availability of medical services to all sectors of the society. Prediction and diagnosis for a particular disease has been simplified since the advent of technology in medical industries. The information obtained from the patients has been digitalized which are recovered from smaller and wearable diseases, examinations and personal check up by the doctors themselves. The risk of the same disease in future can be predicted based on a prediction analysis from the current levels of progression. The prediction analysis can help the patients to delay the progress of the diseases or completely terminate the chances of side effects. Yet there are some diseases which are difficult to predict or derive an accurate and objective decision during the prediction. Alzheimer's disease is one such disease where objectivity in the prediction becomes a hard target, especially in the early stages of the diseases. AD possess very minute and mild symptoms in the early pathogeny and thus this difficulty raises even for experienced medical practitioners. This difficulty can impose a minimal effect over the patients' life and would cause adverse effects. Thanks to machine learning and deep learning technologies that ensures an effective and near accurate solutions in predicting the progress of the life threatening diseases.

Deep learning solutions provide immense benefits when applied over image processing domains as it extracts more

meaningful features compared to traditional approaches. Various Parts of the images and their associations are clearly depicted better using deep learning methodologies. The links which connects the different regions of the brain will produce a better cognitive illustration of the brain and its potential as a network of neurons and operations. A brain is understood by its blood flow and oxygen level relied signal that is captured on a functional and magnetic Resonance Imaging (fMRI) devices. This fMRI information is gained through imaging the brain continuously in regular intervals for different brain activities at different instances and medications. The electronic health monitoring devices delivers the input as the magnetic resonance images will monitor the brain activities and derive the cognitive ability, progress and regress of brain activities. This proposal will monitor the function and establish a pattern of communication between different time series. The application of deep learning strategies is evidently mentioned in this research article by implementing the ideology for the prediction and diagnosis of Alzheimer's Disease where the brain functions are monitored for extending the observations into meaningful patterns and to derive an automated clinical decision support system.

The causes and symptoms of AD are highly varying from case to case due to unique observations in different patients [4]. The hereditary factor attributes to nearly 70% of the AD cases as it is believed that genetic factor is highly influential in bringing the disease over generations [5]. There are no known or effective treatment plans to control the progress or reverse the effects of the disease and in some cases the treatments lead to various other symptoms [6]. Even today, the disease is detected at a very later to advanced stages where the only treatment attempted is to stop the fast progression rate and to protect the brain from further damage. Cognitive decline is one resulting defect due to AD which would have affected the most of a human beings' motor and neurological functionalities, if detected in a very late stage [7]. Hence the AD has to be detected at an early stage to plan for an efficient diagnosis plan along with the remedial plan for side effects. Pathophysiological processes for diagnosis has to begin long before the commencement of the actual diagnosis plan for dementia treatment plans. This pre-emptive diagnosis plan will be helpful for the therapeutic treatment plan for pre-clinical diagnosis. This treatment is vital for patients who have shown Mild Cognitive Impairment (MCI), as they have higher probabilities of developing AD sooner [8].

AD has two predominant symptoms, being the depletion of cognitive function and loss of memory functions and they are the common symptoms of Senile Dementia. When Senile Dementia onsets, MCI will be the primary observations in the elderly. MCI is a nerve damaging disease that results in memory loss, loss of control over motor organs, and inability to remember the common functions of day to day life. Many neuro images have been advocated by medical practitioners in different applications to analyse and understand the root cause of the disease and how to distinguish the biomarkers to narrow down to MCI or AD. Position Emission Tomography (PET) and fMRI are the predominant neuro-images utilized for these operations. The fMRI observed during the resting stage of an affected brain will be captured as a signal fluctuations of blood oxygen level signals and it will be the effective tool for analysing and understanding the modalities of AD/MCI progress levels. This is helpful since the patients are not performing any tiring tasks and resting completely. Resting stage fMRI approach will assist the system to carefully categorize the different sections of the brain network. These images will be used to categorize the patients with just MCI, those with AD and healthy patients termed to be cognitively normal [10]. Researchers utilized a signal matching theory into an application that segregates the brain images based on fMRI information [9]. PET was also widely deployed in various researches that incremented the classification of brain imaging based on the neural activities. PET, due to its added detail and level of attention paid to the features, in brain images, justified that it is the most accurate AD prediction imaging among all other available technologies for brain network imaging [11].

Numerous other methodologies were proposed by researchers to classify the neural images. Support Vector Machines (SVM) are other notable approaches which are being extensively applied in the same domain [12]. SVMs are the foundational techniques employed in traditional approaches in which a machine learning classifier is utilized for identifying the bio-markers. Till date, many approaches rely on the usage of SVM in one way or the other to utilize the SVM directly. Otherwise the technique uses the SVM at a later stage to classify the neural images [8]. Yet the drawbacks of SVM is that the prediction accuracy is achievable up to 60%, especially when a sample structure is used to derive correlation in a complex brain network. This percentage cannot afford a correct clinical support system or prediction system. Medical Science equipped with computer science and engineering, novel and innovative approaches are proposed to solve the problems of predictions and designed with accuracy improvements. Some researchers designed and implemented an auto-encoder with better precision and accuracy for better classification of MCI/AD [12]. Prodromal stage of AD and MCI require a lot of attention and a deep learning architecture to improve the accuracy in detection or classification [13]. Another method of deep convolutional modelling was proposed to analyse the features with two versions namely the supervised and unsupervised learning [14]. The same encoders were extended to sparse auto-encoders and as 3D convolutional networks along with an algorithm to segregate the disease stages according to every individual [15].

The proposed methodology will encompass a novel technique and format for representing the brain operational network, extraction of relevant information for the e-health environments, correlation of information about the brain network. Then, the time series data matrix will be converted into a coefficient format where the computational estimations are simplified by identifying the different associations

between the various brain regions. The next phase comprises of the design and implementation of an automatic encoder where the patients are classified according to their level of disease and AD/MCI predictions. The features of the brain extracted, analysed for correlations based on operations and thus compute the cognitive capacity of the particular patients. Prevention of AD has been complemented with this approach, as this resulted in 20% increase accuracy when compared to the traditional methodologies. This approach has been a pinnacle in the E-Health systems with ability to predict the disease pre-emptively. All automated systems should be enriched with such functionalities, resulting in faster, more accurate and sophisticated medical solutions. The remaining sections are organized into the following, section II discusses the related techniques, section III discusses the data sources, pre-processing, section III introduces the learning model and the representation, section IV discusses the results and evaluations, section V concludes and direct the future work.

2. BACKGROUND WORK

The human brain will be subjected to numerous changes in its development throughout one's lifetime. These physical changes are vital elements in the clinical tests and potential for the detection of AD its early diagnosis. Numerous hours of research have been carried out on the neuro imaging approaches which tracks the pathological changes in human brain that are associated with AD. Consequently, various numerous machine learning techniques were proposed with classifiers operating on images, to identify the persistence of the disease and to derive an action plan for AD. Such studies have highlighted the predominant changes/differences in the brain structures [16,17] usually visible in hippocampus and entorhinal cortex. These structural changes will be clearly distinguished between the healthy brain and AD affected brain. Moving forward, differences identified in cerebrospinal tissues will distinguish the different behaviours in different AD patients [18-19]. A close association can be found between the affected brain tissues of AD patients will also affect the behaviour of the patient [20]. Degeneration of brain cells will be explicit on images which differentiates the patient characteristics and these are observed by different image capturing approaches and devices (sMRI, fMRI), PET, single photon emission computed tomography (SPECT) and diffusion tensor imaging (DTI) These approaches are implemented in different researches which are to be discussed as follows. The traditional methodology to derive the progress of AD is Magnetic Resonance Imaging based mechanisms which also happened to be an automated diagnostic system. This mechanism has two blocks – one for detection of classifying features and the other for categorizing the detected features from the MRI images. The feature extraction techniques are categorized into three primary categories namely voxel based approaches, patch based approaches and regions of interest based approaches. The voxel based approaches will be acting independently without any bias over the brain structures or previous hypothesis. These approaches will concentrate over the local tissues and its density of the brain such as white matter, grey matter and cerebrospinal fluid. The voxel intensities are measured based on morphometrical basis for the classification purposes. The format of representation in voxel intensities are capable of providing better insights, yet they will undergo overfitting problem as they have limited subjects with literally billions of features. This limitation affects the applicability of the approach in neuroimaging for AD diagnosis process. This problem is countered by reducing the dimensions to the required levels of features. The regional information of the brain and its networks are also not considered in this approach.

The next approach to be discusses is the Patch based approaches which segregates the brain images into smaller sections or patches where the features can be extracted. This type will not require regions of interest or any medical expert opinion for further processing and thus reduce the human intervention. Compared to voxel approaches, patch based approaches will be considering dominant dimensionality features to be captured and thus extract minimal brain changes. These techniques tend to study the entire brain images and identify the disease causing features with better perspectives. Yet, the challenges in this approach is that all patches will not be analysed and from those selected features, the system fails to extract the unique features required for diagnosis.

The next approach is the Regions of Interest Approach, which utilizes the predefined brain regions that are structurally or functionally evolved. These regions will be identified based on distinguished features of every area. ROI based approaches will function based on hypothesis derived for abnormal areas of the affected brain. Various techniques have considered the hippocampus region, some utilized the grey matter volume and some utilized the cortical thickness [21, 25, 28, 30]. When the images exhibit low dimensions and cover the entire brain, Regions of Interest technique will be the best technique to be implemented. Yet the ROI approaches the classification based on a whole region of the brain and will not identify the subtle changes in smaller regions in the brain regions. The structural and functional areas in the brain are prone to diseases and they are highly likely to be transmitted to other healthy regions too. A single region may infect other regions and this the other approaches will not be able to detect the affected regions. Moreover, the ROI approaches will demand the advice of medical practitioners and experienced human monitoring. Numerous research works focussed on implementing machine learning techniques on MRI information for AD diagnosis and prediction. Support Vector Machines (SVM), Logistic Regression, Sparse Representation Classifiers, Random Forest based classifications are some significant approaches to be used in AD diagnosis. Linear SVM based implementations[30] work on Ti weightages for AD detection. Variations in the methods along with dimensionality Reduction [30] was implemented on structural MRI data fir AD detection. In both these techniques, SVM binary and multiclass classifiers were

utilized for the detection purposes. An additional classifier was implemented along with other two, demographic and genotype information was utilized to categorize the patients into healthy and affected patients accordingly. A multimodal classification approach [30] implemented a random forest classifier on MRI and PET images. Automated feature selection was designed and implemented for improving the efficiency of the techniques using Hierarchical AdaBoost, SVM and some predefined classifiers. Yet due to the diverse range of classifiers and multiclass unique features, the training model using automated features was difficult to attain and resulted in sub-standard performance.

In the recent past, many deep learning methodologies have outperformed the other mechanisms die to their ability to extract features from the images. The Deep Learning methodologies will require a detailed layer of based hierarchical topology to understand the abstract feature lists observed from the data input. These methodologies impose on learning about easy, simple and low level features from the images which later evolve into high level features obtained through a hierarchical manner. This approach has considerably improvised in the core areas of visual object recognition, recognition of human actions, natural language processing, tracking selected objects, restoration of images, de-noising processes, image segmentation, audio classification and brain computer interactions. Segmentation processes such as detection, registration and classification were achieved better with the help of deep learning models. The latency and ability of extracting hidden features from input images with the help of deep learning models has improved quite a lot. Hence the deep learning models are preferred for implementation widely by many researchers. An auto encoder was modelled based on sparse information for classifying AD/MCI and healthy control. The same system was extended with 3D convolutional neural network and the upgraded model was employed in AD diagnosis. A 2D CNN was also proposed to ease the computations yet promising the same performance and efficiency. Other authors proposed and built a deep belief network where manifold learning models were implemented with both supervised and unsupervised models. Many authors still advocated the 3D CNN models along with deep learning strategies. These deep learning models were enriched with modal stacked encoding networks with a zero masking strategy. This model intended to limit the loss of information due to various reasons. The classification from the unprotected images were extracted with the help of SVM from MRI/PET images. Several SVM classifiers were used to classify the disease in many other deep learning methodologies.

AD is detected after the treatment plan is active, wider plan as the diagnosis reduces the effects of the disease with respect to cognitive ability. The treatment plan will never stop or make the patients healthier but just mitigate the progress of the disease. Early diagnosis of AD will definitely assist the medical practitioners to control the disease. The problems persisting to the AD detections and monitoring the classification of healthy and AD affected patients. Yet a detection scheme should be able to distinguish among various stages of AD itself, let alone the classification of healthy and AD affected patients. In the previous methodologies, we designed a model for detection of affected regions of the brain with image processing techniques. In this paper, the methodology proposed in this article is improvised with deep learning encoders which exhibited better results and performance where the Alzheimer's Disease Neuro Imaging Association datasets are used.

2.1 Pre-Processing of the Data Sources

The study is conducted over the Alzheimer's Disease Neuro-Imaging Association database images which maintains a strong and diverse repository of neuro images of AD affected patients from 2003. The ADNI composes of MRI, fMRI, PET and other classifiers as biological markers. These classifiers and notable features will explain the structural and functional aspects of the brain that leads to detection of AD/MCI. This study encompasses 48 MCI affected patients, 52 normal people to distinguish the healthy and AD affected patients. These patients have consented for being tested and their brain networks are captured through neural imaging devices and genetic samples. Patients age from 63 years to 88 years and patients affected by MCI are aged from 67 to 88 years. Both the test groups belong to the same age group and gender to ensure that no additional parameters decide over the changes in feature extraction and classification.

The fMRI images are downloaded from ADNI in the DICOM format in which it is originally stored. The fMRI images were scanned by a 30 Tesla medical model, the flip angle was 80 degrees, repetition time is 3000 ms, echo time is 30 ms, pixel size is 3.3X3.3 and the matrix size is 64X63. Further information is documented on the ADNI website for reference. Data Processing and Analysis of Brain Imaging (DPABI) toolbox is utilized to analyse the fMRI information. DPABI is a toolkit used for analysing fMRI DATA processing available on a statistical Parameter Mapping Software and REST toolkit. Initially the 10 MRI images that denote the time series with unstable information is removed from the system to ensure that all systems are stabilized. The next set of 130 neuroimages are saved to be utilized in the proposed model. The next smoothening and finally uniformity is achieved. The next step is to pre-process the segmentation process according to the time series. Then, the information is filtered to match the frequency range of 0.01 to 0.08Hz, where the offset is low and the noise is removed. Noise in neuro images usually fall during high frequencies. The next process is to execute the Automated Anatomical Labelled model, which in turn drives the Regions of Interest over the brain network. The entire brain is divided into 90 regions and left/right hemispheres are further divided into 90 regions respectively. Every time series present in the ROI will be calculated with the approximate value of volume in the corresponding ROI. Once the fMRI images are scanned for every patient, a matrix is derived with 90 rows and 130 columns for every region of the brain collectively with respect to an instance of the time series.

2.2 Learning Model and its Representation

The preliminary process of classification is to categorize the images according to the effects of MCI/AD. This process is depicted in Figure 1. The functional connectivity of the brain is evaluated with Pearson's Correlation Coefficient in usual terms. Pearson Correlation Coefficient is represented by the following equation where two variables X and Y are related linearly using the following representation.

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{1}^{n}(x_i-\bar{x})^2 \sum_{1}^{n}(y-\bar{y})^2}} \quad \rightarrow (1)$$

In this equation Xi indicates the ith sample point in the whole sequence of X. Yi , similarly, indicates the ith sample in the whole sequence Y. $\bar{X}$ indicates the aggregate value of all X sequence points and $\bar{Y}$ indicates the aggregate value of all Y sequence points. The correlation coefficient matrix M(90X90) for all 90 regions of interest of the brain. From the obtained matrix, upper triangular matrix will be considered for the research as this matrix is considered to be symmetrical. M(i, j) will illustrate the correlation coefficient between the ith and jth coordinate in the ith and jth region respectively. This original fMRI information will be converted into a format that is inputted into the automatic encoder.

3. PROPOSED METHODOLOGY

3.1 Automatic Encoder using Deep Learning

A network is defined with n layers, for which the pre-learning duties are assigned and separated. For every layer in this network, the following three processes are executed for every layer to be the part of the final network. The three processes should be satisfied before they are employed for the learning process. The weightages of every layer are defined as (W, b) which should be a near zero value. These weightages and parameters must be optimized to reduce the cost of computations. Finally, the third process is to estimate the activation vector that will be considered as the input of the next layer. When the layer *li* is equal to the total number of layers, it will be saved as the input of Softmax Classifier model. The following sections discusses the techniques in detail.

i. Defining the Initial Parameters

The initialized parameters should be optimal in order to obtain the performance enriched design and implementation. The parameters are computed to form the minimized cost function J_{sparse} (W, b). When the best result has to be obtained, best set of weights parameters should be defined by the training model. All the parameters should be some assigned with some values which are very close to zero and it should be negligible. They cannot be easily assigned with any zeros as all layers will perform the same function without any optimizations.

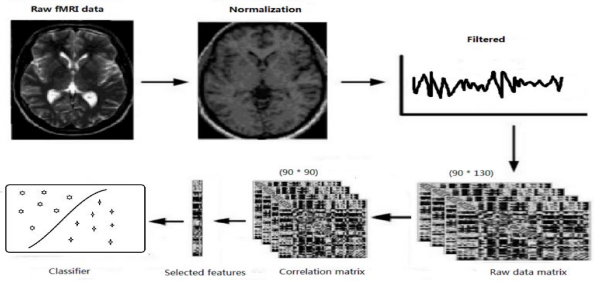

Figure 1: Classification of AD/MCI

ii. Optimization of the Parameters

The algorithm implemented in this research article is BFGS algorithm also known to have a demand for limited memory resources. The algorithm is designed and implemented on the MATLAB environment which has a specific toolkit to achieve the quasi Newton methodology. It is understood that a Newton method would require the partial derivative and inverse Hessian Matrix costing the target matrix with more energy and time. The Newton method cannot be directly implemented in the target function and expect optimal results. Hence the quasi methodology is implemented with an approximate and inverse Hessian Matrix. Bloyd-Fletcher-GoldFab-Shanno Algorithm was improvised into L-BFGS algorithm in order to reduce the consumption of memory resources compared to the original algorithm. The complete set of computations is reduced into the utilization of least vectors for calculating the inverse matrix and hence the final matrix is represented.

For a given f(x) that is twice-differentiable:

1. Choose a starting point x_o
2. Calculate search direction by estimating H^{-1} (method varies)
3. Calculate change in x using the following equation:

$$x^{k+1} = x^k - [H^{-1}]_k * grad(x^k)$$

4. Determine new x value, x^1
5. Determine if method has converged using convergence criteria (gradient)
6. Repeat from step 2 if not optimum

Mean Square Error (MSE), Weight decay and Sparse Punishment are the primary factors that determine the cost function J_{sparse} (W, b). The Mean Square Error is the first element that describes the change in the actual values rather than the expected values. The training set comprises of M samples where (x^1,y^1) and (x^m,y^m) are the sample set. (x^i,y^i) denotes the input in the ith sequence (x,y) represents the label for the sample set. This neural network is improved by implementing the L-BFGS algorithm. The cost function is expressed as the following equation 2. Mean Square Error will be expressed as the following equation 3 for 'm' training samples.

$$J(W, b: x,y) = 0.5 * || hw, b(x^i) - y^i ||^2 \quad \rightarrow (2)$$

$$J_{sparse}(W, b) = \frac{1}{m}\sum_1^m \frac{1}{2} || hw, b(x^i) - y^i ||^2 \quad \rightarrow (3)$$

The next element of the cost function is the weigh decay otherwise known as canonical element. This factor is used to reduce the diverse range of weights in order to reduce the risk of fitting. Higher the weights, higher the weight decay in the research proposal. The range of weights will be reduced which in turn mitigates the output variance and this improves the efficiency of the proposed algorithm. The weight decay parameter λ depicts the importance of weight decay factor in the target function. Small values of λ may result in overfitting and the opposite results in under fitting. Both these cases will not lead to efficiency and will lead to positives. The search objectives implemented in this technique is Mesh Search that reduces over and under fitting problems. The target function with the mentioned parameters is represented as the following equation 4.

$$J_{sparse}(W, b) = \frac{1}{m}\sum_1^m \frac{1}{2} || hw, b(x^i) - y^i ||^2 + \frac{\lambda}{2}\sum_1^{n-1}\sum_1^{S_1}\sum_1^{S_1+1} Wj * Wj \quad \rightarrow (4)$$

The third factor which contributes to the cost function is known as the Sparse Punishment. The brain activity and its sparsity has been discussed in numerous recent studies. Sparsity is defined as the region of interest which will be active at any given time or based on the operation. The operations captured by the fMRI images are sparse as well. Without the estimation of number of hidden layers required, the automatic encoder has to be confined with the least number of hidden layers and they are implemented to identify the characteristics of input images. Given that $a_j^2 = x^i$ indicating the output of jth hidden layer when i is the given input. $\hat{p}$ indicates the average activated value when all input units are considered. This is represented by the following equation where $\hat{p}$ represents the aggregate value for active requests in the jth hidden unit in all m training sample set, p is the sparse parameter in the network. The sparse parameter is assigned to a near zero value and ensure that the active regions are sparse in the entire network.

$$\hat{p} = \frac{1}{m}\sum_1^m \frac{1}{2} a_j^2(xi) \quad \rightarrow (5)$$

The limit of the summation is set to p, when p=0.1, majority of the hidden units will be in an activated state and the sparse parameters will be 0.1. Now the target function will be represented as the following equation 6. Back propagation is the implemented algorithm to track the set of functions and compare the obtained output over the expected output. This algorithm defines the gradient vector of the proposed target function to provide a supervised learning facility. The gradient of the cost function is expressed as the following equation. The error in the last layer will be computed, propagated in the reverse order to address the errors in all layers in prior to the current layer.

$$J_{sparse}(W, b) = \frac{1}{m}\sum_1^m \frac{1}{2} || hw, b(x^i) - y^i ||^2 + \frac{\lambda}{2}\sum_1^{n-1}\sum_1^{S_1}\sum_1^{S_1+1} Wj * Wj + \frac{1}{m}\sum_1^m \frac{1}{2} a_j^2(xi) \quad \rightarrow (6)$$

3.2 Estimation of the Activated Vector

The target function of the proposed methodology is minimized in every layer, the algorithm will compute and deliver optimal parameters and these values are utilized to propagate to the following layers to achieve maximum optimization. When a deep learning automatic encoding network is built, output delivered in the previous layer will be the input to the next layers which identifies the function of whole automatic encoding network. Once all the n layers of the network get through with the learning process and gets trained, the actual process will continue to work progressively. Finally, the activated value of the last layer is generated and assigned as a^{n1}.

i. Defining the Process of Softmax Classifiers

Once all the parameters are defined and the layers getting pre-trained within the network, a proper value will be obtained within the parameter space of the training sets without any labels. With the pre-training results itself, the proposed model would have defined the final solution and this is obtained by the random assignment of values to the parameters equivalent to non-zeros. The softmax regression is also derived that is better than the logistic regression when multiple class categorization is required. The cost function is thus derived to be the following equation.

$$\nabla \square j\, J(\square) = \frac{I}{M}\sum_1^M \{x^I (y^i = j) - p(j|x^i:\square))\} + \lambda \square$$

L-BFGS algorithm utilized in this proposal will attempt to optimize the functions and reduce the cost of the target function. Now the cost function, automatic encoder and the cost gradients are generated to derive the last automatic encoder and deliver it as an input to the Softmax Classifier. This is depicted in the following Figure 2.

ii. Refining the process of Automatic Encoder

The process illustrated in Figure 2 can further be defined by repeating the process defined. Preliminary training is offered to the layers for enriching the knowledge of various characteristics in different regions of interest and different classes. This refinement process is to reduce the cost of entire operations, very similar to the optimization of L-BFGS algorithm.

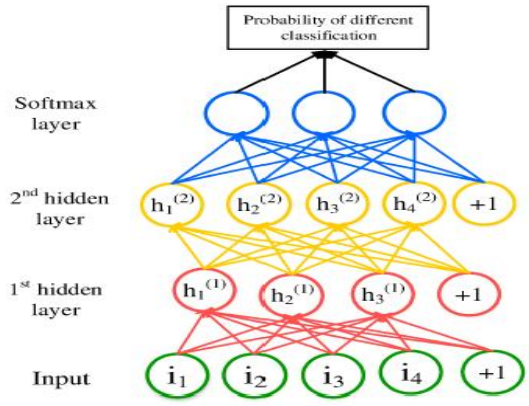

Figure 2: Automatic Encoding Network with 2 hidden layers and classifier layer

3.3 Experimental Results and Discussions

In the performed evaluations, the time series information in the form of a 90X130 matrix (90 being the number of rows and 130 being the number of columns) and the correlation coefficient information in the form of 90X90 matrix are given as input into the machine learning model to train the network. The proposed methodology was compared against SVM and Logistic Regression to understand and analyse the functions of the brain. SVM and Logistic Regression were the conventional model to study the disease prediction models and the deep learning model was implemented in the proposed automatic encoding model. The accuracy of the model differed in varying number of iterations of training and classification, where the machine learning model was tested 10 times with 10 different classifications techniques. The average accuracy level of the 10 accurate values in each computation gives us the integral value of the proposed machine learning model.

Figure 3 illustrates the classification accuracy of three algorithms with respect to the different training time required. The common case is the time series information that is given as input to all the models. From the images, accuracy rate of the automatic encoder is higher than SVM and logistic regression model indicatively. The test results exhibited an accuracy rate of 68% which is 8% higher than SVM and logistic regression models respectively. The accuracy rate of proposed machine learning model is always higher than the other two models. This evidently shows than the new automatic encoder model is much efficient and accurate than classical approaches.

Figure 3: Correlation Coefficient Information

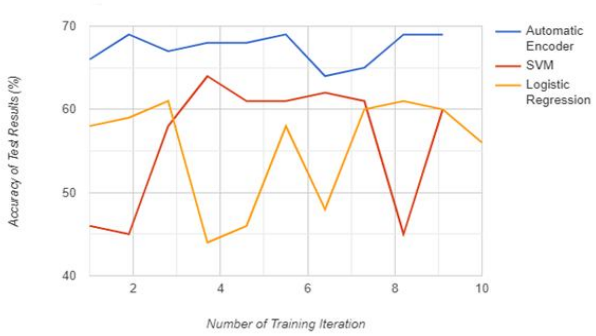

Figure 4: Time Series Information

The input information is in sync with the correlation coefficient matrix, the test accuracies of every model is shown in Figure 4. The input for comparisons is the time series information. The figure depicts the variations in the input values according to the time series information along with correlation coefficient information and that has shown a significant improvement of 20% accuracy when compared to that of other models. The correlation coefficient data will represent the brain activity more in detail and the same will be helpful in predicting the AD. The proposed methodology for generating and classification of correlation coefficient information made considerable improvements and this will support the medical practitioners to automate the decision support systems for enhanced diagnosis plans. The following table 1 will exhibit the accuracy of every model in both the input conditions.

Table 1: Accuracy of all the models with input conditions

Prediction Model	Time Series Information	Correlation Coefficient Information
Automatic Encoder	68%	88%
SVM	62.90%	84%
Logistic Regression	60%	84%

Instead of a time series information, the training model is inputted with correlated information and the accuracy rate of SVM becomes 84%, accuracy rate of Logistic Regression becomes 84% and the proposed model derived an accuracy of 88%. The correlated information has improvised the performance by 20% better than time series information. The connectivity between different regions is fruitful than independent functioning and people's cognitive thinking ability. The test accuracy is better than both the models considering the input variables to produce a better decision support system.

4. CONCLUSION

The deep learning model would require maximum information from various databases and the proposed work should be tested with a diverse range of databases. The information available in the tested data set consists of 100 unique patients' information for assuring the accuracy of the model. Collection of information from wearable devices will enrich the dataset and enable the functions of electronic health monitoring systems. All such information should be stored as credible information onto a Big Data platform to extend the functionality. Models with a bigger scope ensures that prediction of AD and other diseases can be enhanced. The selected list of parameters can be extended with Bayes Optimization to test the model for the same accuracy level. A mixed dataset, computing backgrounds and new parameters will optimize the deliverables as an automated decision support system and prediction model for life threatening diseases.

REFERENCES

1. C. Boucher, M. Bubak, I. Altintas, et al., **From molecule to man: Decision support in individualized e-health,** Computer, no. 11, pp. 40-46, 2006. https://doi.org/10.1109/MC.2006.380
2. L. S. Ronga, S. Jayousi, E. D. Re, et al., **TESHEALTH: An integrated satellite/terrestrial system for e-health services,** Communications (ICC), 2012 IEEE international conference on. IEEE, pp. 3286-2890, 2012.
3. J. Liang, T. Sahama, **Online multiple profile manager for eHealth information sharing,** Communications (ICC), 2012 IEEE International Conference on. IEEE, pp. 3461-3465, 2012. https://doi.org/10.1109/ICC.2012.6364565
4. Burns, S. Iliffe (5 February 2009), **Alzheimer's disease,** BMJ 338: b158. doi:10.1136/bmj.b158. PMID 19196745. https://doi.org/10.1136/bmj.b158
5. **About Alzheimer's Disease: Symptoms,** National Institute on Aging, Retrieved 28 December 2011.
6. **Dementia Fact sheet No. 362**, World Health Organization. April 2012, Retrieved 28 November 2014.
7. R. A. Sperling, P. S. Aisen, L. A. Beckett, et al., **Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease,** Alzheimer's & dementia, vol. 7, no. 3, pp. 280-292, 2011.
8. D. H. Ye, K. M. Pohl and C. Davatzikos, **Semi-Supervised Pattern Classification: Application to Structural MRI of Alzheimer's Disease,** Pattern Recognition in NeuroImaging (PRNI), 2011 International Workshop on. IEEE, pp. 1-4, 2011.
9. C. Hu, J. Sepulcre, K. A. Johnson, et al., **Matched signal detection on graphs: Theory and application to brain imaging data classification,**" NeuroImage, vol. 125, pp. 587-600, 2016.
10. G. Chen, B. D. Ward, C. Xie, et al., **Classification of Alzheimer Disease, Mild Cognitive Impairment, and Normal Cognitive Status with Large-Scale Network Analysis Based on Resting-State Functional MRImaging,** Radiology, vol. 259, no. 1, pp. 213-221, 2011.
11. D. Silverman, **PET in the Evaluation of Alzheimer's Disease and Related Disorders,** New York: Springer, 2009.
12. H. I. Suk, D. Shen, **Deep Learning-Based Feature Representation for AD/MCI Classification,** Medical Image Computing and Computer- Assisted Intervention, MICCAI 2013, Springer Berlin Heidelberg, pp. 583-590, 2013. https://doi.org/10.1007/978-3-642-40763-5_72
13. S. Liu, S. Liu and R. Kikinis, **Early diagnosis of Alzheimer's disease with deep learning,** Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on. IEEE, pp. 1015-1018, 2014. https://doi.org/10.1109/ISBI.2014.6868045
14. F. Liu, C. Shen**, Learning Deep Convolutional Features for MRI Based Alzheimer's Disease Classification,** arXiv preprint arXiv:1404.3366, 2014.
15. A. Payan, G. Montana, **Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks,** arXiv preprint arXiv: 1502.02506, 2015.
16. Y. Chao-Gan, **DPABI: a toolbox for Data Processing & Analysis of Brain Imaging,** http://rfmri.org/dpabi, 2014.
17. de Vos F, Schouten TM, Hafkemeijer A, Dopper EG, van Swieten JC, de Rooij M, van der Grond J, Rombouts SA, **,** 37(5):1920–1929, 2016.
18. Fletcher E, Villeneuve S, Maillard P, Harvey D, Reed B, Jagust W, DeCarli C, **β-amyloid, hippocampal atrophy and their relation to longitudinal brain change in cognitively normal individuals. Neurobiol Aging,** 40:173–180, 2016.
19. Serra L, Cercignani M, Mastropasqua C, Torso M, Spanò B, Makovac E, Viola V, Giulietti G, Marra C, Caltagirone C et al, **Longitudinal changes in**

functional brain connectivity predicts conversion to Alzheimer's disease. Journal of Alzheimers Disease, 51(2):377–389, 2016.

20. Ambastha AK, **Neuroanatomical characterization of Alzheimer's disease using deep learning.** National University of Singapore, Singapore, 2015.
21. Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehéricy S, Habert MO, Chupin M, Benali H, Colliot O, Initiative ADN et al, **Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database.** Neuroimage 56(2):766–781, 2011. https://doi.org/10.1016/j.neuroimage.2010.06.013
22. Anil kumar B, Dr. P. Rajesh Kumar, **Tumor Classification using Block wise fine tuning and Transfer learning of Deep Neural Network and KNN classifier on MR Brain Images,** International Journal of Emerging Trends in Engineering Research, Vol 8, No. 2, 574-583, 2020. https://doi.org/10.30534/ijeter/2020/48822020
23. Suk HI, Lee SW, Shen D, Initiative ADN et al, **Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis**. NeuroImage 101:569–582, 2014.
24. Wee CY, Yap PT, Li W, Denny K, Browndyke JN, Potter GG, Welsh- Bohmer KA, Wang L, Shen D, **Enriched white matter connectivity networks for accurate identification of MCI patients.** Neuroimage 54(3):1812–1822, 2011.
25. Zhang D, Shen D, Initiative ADN et al, **Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers**. PLoS ONE 7(3):e33182, 2012.
26. Zhou L, Wang Y, Li Y, Yap PT, Shen D, ADNI, A.D.N.I. et al. **Hierarchical anatomical brain networks for MCI prediction: revisiting volumetric measures.** PLoS ONE 6(7):e21935, 2011. https://doi.org/10.1371/journal.pone.0021935
27. Greicius MD, Srivastava G, Reiss AL, Menon V **Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI.** Proc Nat Acad Sci USA 101(13):4637–4642, 2004.
28. Suk HI, Wee CY, Shen D, **Discriminative group sparse representation for mild cognitive impairment classification.** In: International workshop on machine learning in medical imaging. Springer, pp 131–138, 2013. https://doi.org/10.1007/978-3-319-02267-3_17
29. Gray KR, Wolz R, Heckemann RA, Aljabar P, Hammers A, Rueckert D, Initiative ADN et al, **Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease.** NeuroImage 60(1):221–229, 2012.
30. Dinu A.J., Ganesan R, Kumar S.S, **Evaluating the performance metrics of different machine learning classifiers by combined feature extraction method in Alzheimer's disease detection.** International Journal of Emerging Trends in Engineering Research, 7(11), 652-658, 2019. https://doi.org/10.30534/ijeter/2019/397112019