ADHD CLASSIFICATION FROM FMRI DATA USING FINE TUNING IN SVM

*ANITHA S¹, Dr S THOMAS GEROGE²

¹PG Student, Department of Biomedical Engineering  
²Associate professor, Department of Biomedical Engineering  
Karunya institute of technology and sciences, Coimbatore, India.  
E-mail : anithaanu43363@gmail.com , thomasgeorge@karunya.edu

ABSTRACT: - ML (Machine learning) is a subset of AI and also improved learning technique has different performance in result over the conventional ML determining the complexity in structures of dimensional data. ADHD is one of the most important neurological disorders and it is represented by different symptoms and we can extract useful the information from FMRI time series. In this paper the ADHD identification and classification is obtained by machine learning techniques. This paper explores an artificial intelligence in unsupervised learning is appropriate to learn features from raw data. The proposed system presented with two stage approaches for ADHD diagnosis which associated SoftMax Regression and SVM fine tuning approach. In the implementation part used FMRI brain images are data sets. The two stage approach shows the high accuracy in performance by using the learning techniques.

Keywords: ADHD Diagnosis, Machine Learning, SVM (support vector machine) with fine tuning, SoftMax Regression, Functional Magnetic Resonance Imaging (FMRI), Artificial intelligence (AI).

I. INTRODUCTION

FMRI has appeared like most significant tool to various regions of brain images with analyzes function. Therefore analyzing and developing functional network of brain images are based on fMRI. Moreover, the analysis of neurological data, the data derived from MRI, positron emission tomography, functional MRI (fMRI). The cognitive progressive impairment of ADHD characterized and evaluated the steady progression of dementia, the early detection and assessment of ADHD is prodromal stage, mild cognitive impairment (MCI). For diagnosis process required valid brain images, the analysis of neuroimaging assisted by diagnostic method which is also probably increase the performance accuracy. Alzheimer disease (AD) is leads to neurological illness represented by memory issues, where the difficulties in word finding and thinking process and decrease the brain volume [1].

AI, deep learning, and neural networks are ML techniques generally used control and solve the real time problems. The machine learning techniques are one of the important applications in an artificial intelligence that is used to learn the raw data, it is used to interrelating regions and fundamentals, then it is applied for diagnose the neurological disorder. The neurological disorder prediction has implemented with machine learning algorithms, but in this paper framework we used SVM with fine tuning and along with SoftMax regression model features, are evaluated the classification and overall system
performance accuracy. The rest of this paper is organized as follows: section II presents related works. Section III is proposed model. Section IV implementation and section V is results and discussion Section VI is conclusion.

II. RELATED WORKS

Here we discussed about related works of ADHD, An important possibility was analyzed and examined as mean of the COSLOF fundamentals are brain region between the feasible pair of voxel representation time [1].

The classifier is presented to extract connectivity features of discriminative function from healthy controls of patients. Three level processes are involved in classifier schizophrenic patients and healthy status, that three levels are feature selection, reduce dimensionality by LLE, and C-Means clustering approach for data classification [2]. Many reviews have been occurred ADHD, the work expand the consideration of projection and extraction of image feature approach on this form of data, combined like singular value decompositions, random forest [3]. In addition, multicenter dataset framed by ADHA-200 consortium respects trst approach for every determination dependent in MRI-noticeable primary characteristic and practical interconnection yield though rs-fMRI. The import of diagnosis approach implemented by predictive power triple sets of attributes: non-imaging, network feature and structural data [4]. An enhanced subjective-adaptive approach was combined SICE networks for characterized by classification. The joined weight adaptively learns for each worry with adaptability and SICE portrayal by and large prompts set of organizations alongside specific sparsity levels. Every single level is captured by specific patterns of connectivity, which is different from other sparsity levels. The work discuss that the SICE networks at various level of sparsity may not required be linear across all the subjects. It has some considerations following: same class label sharing, demonstration of brain networks with various subjects’ age, gender, etc [5]. ADHD classification is determined by fMRI data to participant in diagnosis process and before comparing the dimension reduction schemes used principle component investigation system to limit the element vector measurement [6]. AI grouping techniques referenced above for ADHD rely upon single enhancement strategy. At any rate one hyper-limit ought to be picked before setting up a classifier. This regularly prompts tedious instructional meetings and grouping failure with changes in model size. For example, while using the SVM to describe smooth mental shortcoming subtypes [7], Haller [8] researched the gamma limit iteratively from 0.01 to 0.09. The rule reason is that the upgrade issue in planning system commonly incorporates more than one objective (for instance, two regular goals in SVMs are amplifying the edge of partition and limiting experimental mistake). Nonetheless, conventional AI strategies by and large utilize a compromise boundary to summarize all the destinations into one furthermore, accordingly transform a multi-target enhancement issue into a solitary target one. Consequently to acquire a productive classifier ordinarily preliminary what's more, mistake measure is required for picking reasonable boundaries. The ADHD-200 is a freely accessible multisite information base which has been generally utilized for looking at the aftereffects of various grouping strategies. Utilizing just the data of the fMRI information accomplished 61% order precision in [9], while by considering phenotypic data, for example, sex, age, handedness, verbal, and IQ, the presentation expanded to 62.5% in [10]. As of late, there has been an extraordinary propensity for utilizing distinctive example acknowledgment techniques in entire cerebrum neuroimaging information, for example, fMRI information, to
group mental sicknesses. In [11], the neighborhood anatomical ascribes were considered as highlights. The point of the technique was relationship calculation between various dynamic territories in the cerebrum and to uncover them as diagrams. The creators found that the highlights dependent on meager converse covariance are all the more remarkable in the classification. Also, by thinking about sexual orientation as phenotypic data, the arrangement exactness expanded up to 80%. The work in [12] used a subject flexible strategy to join meager reverse covariance evaluations (SICEs) for showing of cerebrum accessibility associations. The precision of their work was 72.5% among ADHDS and routinely made controls (TDCs). In [13], features were remove from both rest state fMRI (R-fMRI) information and primary MRI information by processing the worldwide found the middle value of practical network maps past the whole mind. Discriminant highlights were chosen utilizing an inadequate component choice calculation. The exactness speed of the procedure was 76.19% and 92.86% for multiclass and two class groupings, separately. The creators in [14] investigated the social also, social factors in multisite ADHD-200 data and proposed the multisite approach reliant on the decision made tress. Their methodology for grouping ADHD subtypes and TDCs arrived at around 70% order precision.

In [15], the markers proposed a bi-target ADHD gathering plan subject to the L1-standard SVM by contemplating the test bumble and the edge of segment as two objections. The request results for Kennedy Krieger Institute (KKI), Peking University 1 (Pek 1), and Pek 1-3 are 81.25%, 86.67% and 81.08% exclusively. The creators in [16] introduced another strategy for highlight choice. Their chart piece based primary component choice (gk-SFS) technique was intended to locate the nearby to worldwide underlying data rather than just vector-based features. A learning framework subjects to gk-SFS was used to examine the ADHD contamination. The procedure achieved 63% gathering precision for New York University (NYU) site (118 ADHDS and 98 TDCs). In [17], a profound change technique (DTM) was created which abuse the discriminate torpid component space and checks it in the accompanying layer for portrayal of 465 rights-gave male subjects. The softmax authorization work was used in the decision layer to portray f-MRI data which came to 70.36% exactness.

III. PROPOSED MODEL

In this paper we proposed softmax regression and SVM with fine tuning methods and overall system given in Fig.1. Model based on dual stage learning approach. SoftMax regression is first stage learning is used to label the healthy and unhealthy data and SVM is second stage learning with fine tuning is used to extract features of brain images, the layers are called as double layer network. Learn the significant features are extracted by neural network make it reasonable for processing the more signals in monitoring condition. As per the proposed method classification achieved from brain data and the implementation results are highly efficient than conventional methods.
Fig. 1 Proposed Model

Above shown figure represents two stage learning method for ADHD diagnosis. The softmax regression is used to learn the features of health status and the features are extracted by using SVM. Some more learned highlights of pixels acquired by neighborhood highlights, which is classified.

3.1 SOFTMAX REGRESSION

In neural organization, the last layer softmax relapse is evaluated to manage the multiple classes’ classification. Its implementation is very easy and normal. Assume the preparing set of \( \{x_i\} \) \( M_i = 1 \) alongside a named set of \( \{y_i\} \) \( M_i = 1 \), where, \( x_i \in \mathbb{R}^{N \times 1} \) and \( y_i \in \{1, 2, \ldots, K\} \). For each info set \( x_i \), the model attempts to appraise the probability of \( p(y_i = k|x_i) \) for every one of the marks of \( k = 1, 2, \ldots, K \). Thus the softmax relapse creates a nonexistent vector that gives the K-assessed probabilities giving information test of \( x_i \) having a place with each name. The presumption of \( \mathbf{h}_\theta \) is as per the following:

\[
\mathbf{h}_\theta(x_i) = \begin{bmatrix}
p(y_i = 1|x_i; \theta) \\
p(y_i = 2|x_i; \theta) \\
\vdots \\
p(y_i = K|x_i; \theta)
\end{bmatrix} \frac{1}{\sum_{k=1}^{K} e^{\theta_k^T x_i}} \begin{bmatrix} e^{\theta_1^T x_i} \\
e^{\theta_2^T x_i} \\
\vdots \\
e^{\theta_K^T x_i}
\end{bmatrix}
\]

Where, \( \theta = [\theta_1, \theta_2, \ldots, \theta_K]^T \) is the softmax relapse model boundary. It should be noticed that \( k = 1 \) \( K e^T \theta \) \( x_i \) standardizes the appropriation, so the presumption amount of the parts are
approaches 1. To limiting the expense capacity of $J(\theta)$ done by supposition model.

$$J(\theta) = -\frac{1}{M} \sum_{m=1}^{M} \sum_{k=1}^{K} \left\{ y^m = k \right\} \log \frac{\hat{y}^m_{\theta_{k}}}{\sum_{k=1}^{K} \hat{y}^m_{\theta_{k}}} + \frac{\lambda}{2} \sum_{l=1}^{L} \sum_{j=1}^{q} \theta_{l,j}^2$$

In Softmax regression parameters are close to zero, because model parameter minimizes the cost of function.

To carefully make raised of both term of weight reduction (for each $\lambda > 0$), cost capacity of $J(\theta)$ and check that softmax relapse model hypothetically has a flexible arrangement. Moreover, softmax relapse has flexible answer for the issue of characterization, which is utilized to decide the probability that an example will have a place with any wellbeing status name.

![SoftMax Regression](Image)

**Fig.2 SoftMax Regression**

### 3.2 SVM WITH FINE TUNING

SVM is the data classifier which is incorporated with linear and non-linear model and this is significant with classification and prediction. In our proposed system used linear kernel...
model, this model classifies individual weight of data features. Each node’s significance in linear kernel SVM, each features are opposed to each other significance.

The importance is straightforwardly corresponding to amount of the loads of association from the hub.

3.2.1 PARAMETER TUNING

The fine tuning parameters in SVM are implemented for feature space, the tuning process will be selected holdout set parameter model and randomly chosen subset with respective dataset can achieved accuracy of classification and system performance.

For example randomly selected set of 30 subjects was used to tune the SVM leaf disease classification parameter. To use limited number of features <900 for two conditions: conventional concept provided high accuracy of features on the order of 90 for evaluating the relevant network. Grid Search approach was included in implementation process for an interval number of top features and N. The value of N and number of features achieved accuracy were utilized through total dataset.

IV. METHODOLOGY

First stage of implementation based on SoftMax regression, the problem of the parameters evaluated, it is containing “N” number of input and “N” number of yield measurement with inadequate separating and term of weight reduction is $\lambda$ in regression model. An unsupervised learning method is proposed and it is the advantages to enhance the accuracy of diagnosis process, basically the unsupervised learning is depth through the features of network weights, the important of the models are used to fill the gap between extraction and signal processing by using AI techniques.

SVM classification depend correlations between the ROIs. As input of each data accuracy was found 94%. This accuracy shows that our enhanced SVM can be applied to fMRI data with highlight choice and boundary tuning successfully. The executed classifier performance was more accurate. As shown in the below comparison of the proposed and conventional method.

| PARAMETER | EXISTING METHOD | PROPOSED METHOD |
|-----------|-----------------|-----------------|
| ACCURACY  | 86.32%          | 94%             |

Table 1: Accuracy comparison table
4.1 IMPLEMENTATION

The below structures the implementation of our project we have taken a total of 164 fmri data: 87 AD, 77 Normal Control, and Number of Patients: 38 AD, 20 Normal Control. The classifiers as mentioned above we have used Softmax regression, Stacked Autoencoder for training and prediction. For classification we have used SVM with fine tuning. The number of features as follows 1 - 90x130 = 11700, 2 - 90x90 = 8100, 3 - 81x45 = 4005, 4 - 41x21 = 861. The dataset is structured as below,

Dataset 1: The original data, 80 different brain areas for 80 rows, 120 columns allocated for 120 samples at different event.

Dataset 2: Figure the connection network from the dataset 1, Dij has a place with [-1, 1], addresses the relationship coefficient between i-th and j-th zone.

Dataset 3: Since the relationship network is a symmetric matrix.

Dataset 4: Select 45 sub-zone in entire brain network to reduce the input’s measurement. The accuracy we get after the implementation is 94% using the fine tuning in SVM.

Fig 3: Accuracy comparison Graph
Fig 4: Matlab GUI

The above figure shows the GUI created to run the code with ease.

Fig 5: Accuracy Value for Single Dataset

The Fig 5: shows the accuracy values 100% for a single dataset during the execution process.

V. RESULTS AND DISCUSSION

The proposed SVM model has been subject to all F-MRI signals. The machine learning was developed in the MATLAB language using softmax regression and SVM with fine tuning method. The proposed SVM model yielded an accuracy of 94% on the ADHD-200 holdout data. The presented model is based on the promising outcomes obtained. Our
outcome shows that it is feasible to diagnosing ADHD subjects. The below graph shown the accuracy value.

Fig 6: Output Graph

VI. CONCLUSION

This paper has investigated fMRI data based ADHD diagnosis that there may be the inherent, learning features from raw data is important to choose proper label using SoftMax regression and SVM with fine tuning to data feature extraction together as a combined approach for an exact ADHD identification process. The ADHD early detection and classification is really helps to treat the affected region and keeping away from the long term impacts. In this paper, we focused on deep learning with hidden features representing ADHD data for diagnosis. To this end, considered that deep learning able to evaluated the imaging data and we implemented two stage approaches for diagnosing Attention Deficit Hyperactive Disorder disease for the first time.

REFERENCES

1. Li, Shi-Jiang, Zhu Li, Gaohong Wu, Mei-Jie Zhang, Malgorzata Franczak, and Piero G. Antuono. "Alzheimer disease: evaluation of a functional MR imaging index as a marker." Radiology 225, no. 1 (2002): 253-259.
2. Shen, Hui, Lubin Wang, Yadong Liu, and Dewen Hu. "Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI." Neuroimage 49, no. 4 (2010): 3110-3121.
3. Eloyan, Ani, John Muschelli, Mary Beth Nebel, Han Liu, Fang Han, Tuo Zhao, Anita D. Barber et al. "Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging." Frontiers in systems neuroscience 6 (2012): 61.
4. Bohland, Jason W., Sara Saperstein, Francisco Pereira, Jérémy Rapin, and Leo Grady. "Network, anatomical, and non-imaging measures for the prediction of ADHD diagnosis in individual subjects." Frontiers in systems neuroscience 6 (2012): 78.
5. Dang, Shilpa, Santanu Chaudhury, Brejesh Lall, and Prasun Kumar Roy. "Learning effective connectivity from fMRI using autoregressive hidden Markov model with missing data." *Journal of neuroscience methods* 278 (2017): 87-100.

6. Shao, Lizhen, Yadong Xu, and Dongmei Fu. "Classification of ADHD with bi-objective optimization." *Journal of biomedical informatics* 84 (2018): 164-170.

7. Cao, Jin, Li Zhang, Bangjun Wang, Fanzhang Li, and Jiwen Yang. "A fast gene selection method for multi-cancer classification using multiple support vector data description." *Journal of biomedical informatics* 53 (2015): 381-389.

8. Haller, Sven, Simon Badoud, Duy Nguyen, Valentina Garibotto, K. O. Lovblad, and P. R. Burkhard. "Individual detection of patients with Parkinson disease using support vector machine analysis of diffusion tensor imaging data: initial results." *American Journal of Neuroradiology* 33, no. 11 (2012): 2123-2128.

9. Eloyan, Ani, John Muschelli, Mary Beth Nebel, Han Liu, Fang Han, Tuo Zhao, Anita D. Barber et al. "Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging." *Frontiers in systems neuroscience* 6 (2012): 61.

10. Brown, Matthew RG, Gagan S. Sidhu, Russell Greiner, Nasimeh Asgarian, Meysam Bastani, Peter H. Silverstone, Andrew J. Greenshaw, and Serdar M. Dursun. "ADHD-200 Global Competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements." *Frontiers in systems neuroscience* 6 (2012): 69.

11. Bohland, Jason W., Sara Saperstein, Francisco Pereira, Jérémy Rapin, and Leo Grady. "Network, anatomical, and non-imaging measures for the prediction of ADHD diagnosis in individual subjects." *Frontiers in systems neuroscience* 6 (2012): 78.

12. Zhang, Jianjia, Luping Zhou, and Lei Wang. "Subject-adaptive integration of multiple slice brain networks with different sparsity." *Pattern Recognition* 63 (2017): 642-652.

13. Qureshi, Muhammad Naveed Iqbal, Jooyoung Oh, Beomjun Min, Hang Joon Jo, and Boreom Lee. "Multi-modal, multi-measure, and multi-class discrimination of ADHD with hierarchical feature extraction and extreme learning machine using structural and functional MRI." *Frontiers in human neuroscience* 11 (2017): 157.

14. Itani, Sarah, Fabian Lecron, and Philippe Fortemps. "A multi-level classification framework for multi-site medical data: Application to the ADHD-200 collection." *Expert Systems with Applications* 91 (2018): 36-45.

15. Shao, Lizhen, Yadong Xu, and Dongmei Fu. "Classification of ADHD with bi-objective optimization." *Journal of biomedical informatics* 84 (2018): 164-170.

16. Wang, Mi, Biao Jie, Weixin Bian, Xintao Ding, Wen Zhou, Zhengdong Wang, and Mingxia Liu. "Graph-kernel based structured feature selection for brain disease classification using functional connectivity networks." *IEEE Access* 7 (2019): 35001-35011.

17. Sartipi, Shadi, Hashem Kalbkhani, Peyman Ghasemzadeh, and Mahrokh G. Shayesteh. "Stockwell transform of time-series of fMRI data for diagnoses of attention deficit hyperactive disorder." *Applied Soft Computing* 86 (2020): 105905.