ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse161942020.pdf
https://doi.org/10.30534/ijatcse/2020/161942020

Attention deficit hyperactivity disorder prediction using Machine learning

Adarsh Kushwaha[1], Anmol Mishra[2], Ayushi Jain[3]
[1]Jaypee Institute Of Information Technology, India, adarshkush1311@gmail.com
[2]Jaypee Institute Of Information Technology, India, anmolmishra.jiit@gmail.com
[3]Jaypee Institute Of Information Technology, India, ayushijain.jiit@gmail.com

ABSTRACT

Attention deficit hyperactivity disorder is a mental disorder and after a survey it can be concluded that 5% children and 3% adults are affected by this condition. It can cause learning disability in children which is not tolerable in this learning phase of life. Deep analysis is required of this disease because it is one of the hardest diseases to be diagnosed. The subtle symptoms of this disease are misunderstood most of the time. Most of the children are not properly diagnosed and always get remark by teacher that your child is not performing in class, he is not giving his best effort. Mental disorder is illness which deteriorate the whole life of a person when neglected. ADHD is the most severe and neglected mental disease found in children and in adults.

We as a society found it very difficult and shameful to adopt the truth that he or she may be suffering from mental illness that is the main reason this kind of disease never get diagnosed and create difficulties for humans in future. In fact diagnosis of this kind of disease is very difficult to identify due to its very subtle symptoms. Experiment results illustrate the technical aid using the FMRI scans of the brain and unsupervised learning with a decent percentage of prediction accuracy.

Key words: Computer Aided detection, Feature Extraction, Machine learning, mental disorder detection, Decision tree

1. INTRODUCTION

Human Brain, spinal Cord and nervous system under one head is known as Central nervous system. Attention Deficit Hyperactivity Disorder [1] is a complex neuro disorder. A person suffering from ADHD has a relatively slow and different brain development process as compared to normal person. In this mental disorder the victim's brain loses the self-controlling powers, and develops a slow responding and understanding ability. ADHD can be transmitted in genes from parent to children. Exposer of pregnant women to chemicals such as Nicotine can also be the reason from ADHD. Premature birth and extremely low weight at the time of birth may lead to ADHD.

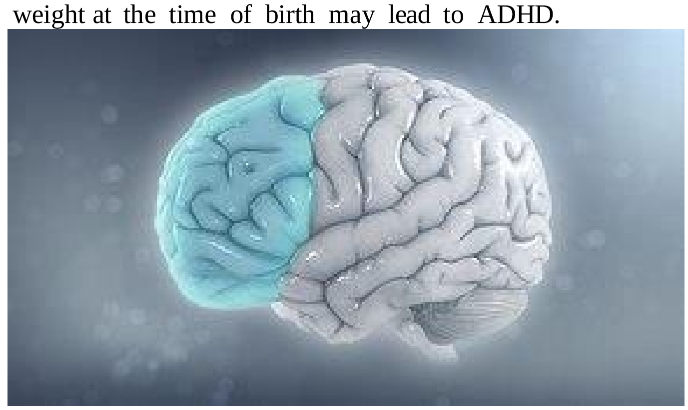

Figure 1: The light blue portion is affected with ADHD

Figure 1 depicts the portion of the brain that is affected by ADHD. This has been proved by the medical science that there is the reduction of volume and thinning of cortex in the left prefrontal of the brain [2]. ADHD brain scans are observed to be of smaller size than normal healthy brain due to slow development and reduced volume. ADHD brains scans are observed to have low blood flow as compared to healthy brain which leads to reduced concentration in performing any activity.

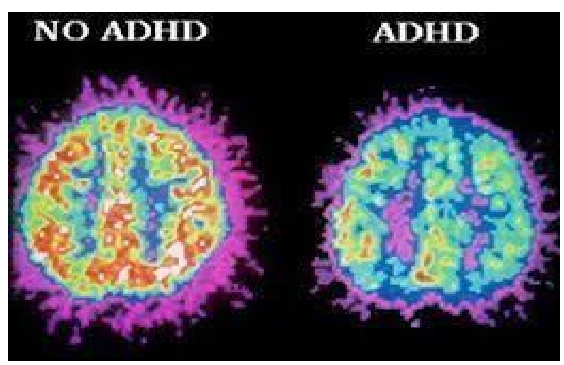

Figure 2: Non-ADHD vs ADHD

Figure 2 represents brain scan of ADHD and Non-ADHD brains. To diagnose the disorder the doctors first observes the patient's habits, talks to their family members about their behavior.

Symptoms of ADHD in human being:
1) Person suffering from ADHD is not able to respond to the other persons efficiently.
2.) ADHD persons develops the self-focused behavior thus they interrupt other people in their activities to gain their attention.
3) Children suffering with ADHD losses their control over patience and troubles waiting for their turn during and class sessions, games our any daily activities.

Magnetic Resonance Imaging (MRI) [3] [4] generates the detailed images of the different organs and tissues with the help of radio waves and magnetic fields. The signal generated are the measurement of the density with which the waves are reflected back to the scanner by the tissue. The doctors will be able to detect the problem in an early stage. With the use of cutting-edge methods[5] [6] and techniques we aim to collaborate to the betterment of the Attention Deficit Hyperactivity disorder Detection and classifications and in turn help the people suffering from mental disorder.

Most of the research works carried out earlier by t/he different researchers make use of the most common Global datasets available namely - ADHD-200 [7], ABIDE. Both dataset include the high definition structural MRI and functional MRI scans of both patient suffering from ADHD and healthy persons. In some of the research works we came across this imaging dataset is combined with Phenotypic data includes age, sex, handedness, Intelligence level, and scanning sites of the brain. Phenotypic data was founded more useful in building models to detect the disorder with much greater accuracy.

Since the imaging data includes high definition images it quit important to decrease to resolutions, dimensionality, and linearity of the dataset. In order to achieve the same almost all the researches have used the most common dimensionality reduction techniques i.e. PCA [8], LDA and some researches includes FFT. In order to achieve higher accuracy some images were rotated at different angles, some were blurred, and some were sharpen [9]. Different methods were used to increase or decrease the noise as per the requirement and training. The feature maps thus obtained were made to undergo several other processes according to the model requirement based on the researcher requirement for example scaling of the dataset, k-fold cross validation, averaging etc.

At this stage the feature map obtained are finally served to machine learning models [10]. All the researches includes SVM [11] model as one of the important model to classify the healthy vs ADHD suffering patient. SVM and its different versions are always considered as the top candidate in classification task because of its higher accuracy and efficiency to differentiate between two or more classes. In different researches different models are combined with SVM as one of the important part to effectively produce the results for example CNN + SVM (linear/kernel).

The other used model were feed forward neural networks i.e. ELM. The results of different research contributions varies from 50% to 91% accuracy. The average accuracy rate is 72% to 76%. The highest accuracy of approx. 91% was achieved with ELM mode.

After studying and analyzing the research paper and the parameter affecting Attention Deficit hyperactivity disorder we had decided to work on 2 parameters namely variations in functional and structural MRI scans, Physical factors such as age, IQ level etc., and thus collected the dataset. After collecting the dataset our first aim was to successfully clean the dataset and preprocess it. We applied several cleaning and pre-processing techniques [12] to achieve the desired output and prevent the model form unwanted noise. Then we perform the correlation analysis to collect the knowledge of the degree of correlation among parameter. This step of finding the correlation among the correlation analysis also helps us to understand the dataset more quickly and clearly before visualizing and analyzing the dataset. Now, at this point of time we divided the dataset into two parts train and test in the ratio of 70:30 and rest all operation except testing is done on training dataset. The next step is to tune the hyper parameters. We had also applied the cross-validation so that our model does not over fit or under fit the model so as to reduce the false negative probability. The last step is to train and predict the different models and compare their results on the parameter of Accuracy, precision and recall.

After successfully training the machine learning models on personal characteristics dataset to move to CNN VGG-16 model. We trained the model on MRI image dataset using predefined weights available globally. We augmented the images by applying filter, cropping and rotating them. Then they are served as input to the VGG16 model.

2. LITERATURE SURVEY

Classification of ADHD and Control Adults with Phase Space Reconstruction of Electroencephalography (EEG)Signals is discussed [17], the features and tools used are Electroencephalography signals, support vector machine, k-nearestneighbour, Naive-Bayes classifier, enhanced probabilistic neural network, NDC and found that the accuracies were 80% for neural dynamic classifier, 76.7% for enhanced probabilistic neural network, 80% for support vector machine, 73.3% for k-nearest neighbourand 76.7% for naive-Bayes classifier. Various statistical features have been constructed upon reconstruction of phase space of signals. There is a high accuracy of 93.3% under the eyes-open, 90% under the eyes-closed, and 100% under the CPT condition.

Dilated 3 dimensionconvolutional neural network method used for classifying 3 dimension MRI scans convoluted with feature maps and gave 76.6% accuracy on structural MRI ADHD images and 82.7% on SZ fMRI EPI imagesof a Schizophrenia dataset, demonstrating better performances than some other state-of-the-art methods.The accuracy achieved in this using single EPI image level.This method could also be used in Computer-Assisted Diagnosis and real-time fMRI analysis [18].

Classification of Structural MRI Images in ADHD using 3D Fractal Dimension Complexity Map [19], fractal dimension map based 3D CNN was used for Automatic diagnosis of ADHD with significant results in representing gray matter density data sparsely. The research is effective in supporting the efficiency of MRI images in ADHD.

In [20], Social media and big data analytics were used in predicting the mental health of an individual by Co-training with and without SVM, RF, NB and classifiers with Co-training outperformed the state of the art models by a margin of 3%.

Detecting ADHD Patients by an EEG-Based Serious Game, an EEG based serious game using RBF Kernel model, EMOTIV data and keyboard data was used for Detection level of ADHD, with 98.4% for RBF Kernel and 64.73% and 64.74% 1 ADHD subject’s data holdout and 2 ADHD subject’s dtaa holdout respectively [21].

In [22] Evaluating the Effectiveness of a Group CBT for Parents of ADHD Children, an experimental study Quasi observed the culturally aware CBT treatment groups’ effect on mental health and quality of life (QoL) of Chinese parents with children with ADHD Effect. Effect of CBT on mental health and improved QoL significantly mediated after reduction in parenting stress and on reduction of dysfunctional attitudes, the effect of CBT on improved QoL significantly mediated.

Event-Related Potentials for Diagnosing Children and Adults with ADHD, there was a study about discriminating participants based on ERP classifiers between ADHD and healthy controls. A total of 49 papers were scanned and selected out of which 7 fulfilled the inclusion criteria and gotselected for final review [23].

ADHD and Psychiatric Comorbidity: Functional Outcomes in a School-Based Sample of Children, Teachers and parents of South Carolina and Oklahoma were selected and went through different screenings using Logistic Regression. Children with ADHD and low academic performance showed an increase in psychiatric disorders and anxiety, ADHD alone associated with school discipline [24].

J – Eros [27] which is a model selection scheme i.e. It picks an optimum value for k from the training data for k – nearest neighbors. As compared to the current algorithms, the model showed an increase in accuracy by 20%. The main here in this research is to be able to design such GPU based parallel computing algorithms that analyze the big fMRI data and also by taking some pairwise relations between each voxels without any constraints. The paper [27] also discusses also presented a strategy that was based on GPU which uses Pearson correlations to allow functional connectivity to be calculated for big fMRI data [27].

In [28] paper a new feature selection method based on Relative Importance and Ensemble Learning (FS_RIEL) was introduced through which high dimensional feature spaces get reduced to more refined subspace. Decision trees are used to calculate the relative importance of features. While maintaining the low dimensions of FCs feature space, a forward – backward selection algorithm is applied on combined features to increase the diversity of new feature space. When they compared conventional feature selection methods with FS_RIEL, it showed that the later increased the ADHD classification by about 15% in both adults and children and gave around 80% - 86% accuracy.

The chance accuracy set by the researchers was 64.2% for the data used at the training time. The logistic regression classifier, linear, quadratic, cubic and kernel support vector machine performed better on all the combinations of the feature sets. The best performance was shown by kernel-rbf SVM. Maximum accuracy obtained was 75% from linear SVM [29].

The paper [30] follows the method of deep learning to diagnose Attention Deficit Hyperactivity Disorder. This paper [30] increases the performance of CNN algorithm by taking the input dataset of three popular types of features of fMRI data. The three features were Regional Homogeneity, Voxel Connectivity and low amplitude of low frequency fluctuation. These were combined to create a 3D dataset, which further results in 3D CNN model. As a result this model has encouraged the diagnosis accuracy of ADHD disease. They have also implemented 10-fold cross

validation which helped them in selecting different sets of validation in each occurrence. Performance of methods using classifiers like multi-kernel learning and SVM were compared to the proposed method using 3D CNN as classifier based on the ADHD–200 dataset.
The proposed method gave the highest accuracy of 65.67% which is supports.

3. SOLUTION APPROACH

We have implemented Ensemble learning methods and deep learning model to predict the ADHD disease through MRI scans and IQ level datasets. Ensemble learning is the stand out technique in the field of ADHD disease prediction and we have attained trustable accuracy percentage on that. Ensemble learning is aggregation of different classifiers to attain better results. Sometimes, the only classifier is not sufficient for classification due to large dataset or different kind of features in a dataset then comes the role of ensemble learning when one classifier is not sufficient and called as “Weak classifier”. Ensemble learning methods benefits are as follows:
1. Easy to train
2.Takes less time
3.Scalable
4.Improved results

Ensemble learning methods:
1.**Bagging:** Bagging Classifier [13] is used to reduce variance of the aggregated model which is received after learning all weak learners. However, it’s very much near to impossible to reduce variance of resultant model due to less size of dataset. So, the idea is to create one or more than one bootstrap sample and learn these samples one by one and then take average of all the samples after learning. Two ways are there to aggregate all the models and find the output class:
a. Hard Voting: Take the output of each samples and class with highest votes is taken as a result.
b.Soft voting: In this method we calculate the probabilities of each class obtained and the take the average of probabilities.
Formula used for classification:

$$sl(\,.\,) = \arg k \max\,[card\,(l \mid (\,.\,) = k)]$$

Above formula is used to take the class with maximum votes.

2. **Gradient Boosting**: Boosting [13] is basically training weak learner models in a sequential way to reduce bias. Hence, model with a higher value of bias is trained at first.

Now talking about Gradient Boosting the model which we build is the weighted some of less powerful learners.
Formula used:

$$sl\,(\cdot) = \sum Ll = 1\; cl \times wl\,(\cdot)$$

where cl is a coefficient and wl are weak learners
It views the whole issue of reducing the bias as a gradient descent. Every time we learn a weak learn we learn it opposite to the value of gradient descent of the recent fitting error of recent ensemble model.
Formula used for Gradient descent process applied on ensemble model:

$$sl\,(\,.\,) = sl{-}1\,(\cdot) - cl \times \nabla sl - 1\mathrm{E}\,(sl - 1)\,(\cdot)$$

In the above formula, Fitting error of model = E ($\cdot$) and cl is a coefficient of step size.

3.**Extra Tree Classifier:** Extra tree classifier [14] is an ensemble machine learning algorithm which works on decision tree. In this model several decision tree estimators are created based on random subsets of the dataset. Since each dataset is based on random subsample mostly the trees are built different from each other. Now at the time of testing, each instance is evaluated on all the trees formed and then the class with maximum votes is returned.

3.1 Deep leaning model

a. CNN VGG 16: VGG-16 [15] is a 16 layered CNN based model that focuses more on convolution layers than other hyper parameters. In VGG-16 each convolution layer uses 3x3 filters with stride of 1 and max-pooling layer uses 2x2 filter with stride of 2. At the end there are three fully connected layer followed by soft-max layer at the output. There are all over 13 convolution layers and 5 max-pooling layers followed by 3 fully connected layers and finally the 1 soft-max layer.

i) Convolution layer: The input in the form of matrix is fed to the convolution layer. In this layers several filter are applied and the output is send to next layer.
ii) Max-pooling layer: The dimensions of input are reduced by taking the maximum value in the stride.
iii) Fully connected layer: In fully connected layer each node is connected to every node of the next layer.
iv) Softmax layer: In this layer the output from the fully connected layer is converted to discrete probability distribution so that they can be classified into classes.

Mathematical modelling of CNN VGG 16

• Error formula:

$$e = \frac{1}{n}\Sigma_k \min(i) d(C_i, C_k)$$

where c is matrix of classes sorted in decreasing order and C is a matrix of truth class.

Loss function:

$$(\hat{y}, y) = -\sum 1000\, yi \ \log(\hat{y}\, i)$$

where y is a vector of prediction and ŷ is output vector

3.2 Dataset

Government introduced a competition of predicting ADHD disease through two types of dataset one is phenotypic dataset which is in number and other is Brain MRI scans of ADHD and non ADHD patients. Different universities provided their datasets. Datasets [16] were mainly collected by public who have released 776 fMRI scans in which anatomical dataset was aggregated using different imaging sites. Hence, we have used 2 types of dataset.

a. Phenotypic dataset: This Phenotypic dataset consist of phenotypic information (personal information) of 221 ADHD and non ADHD people. The phenotypic information contains 8 features that are gender, HI, med, aggression, perfIQ, VerbIQ, fullIQ and AD.

b. Brain MRI Scans: Brain MRI scans dataset consist of 98 brain scans of non ADHD persons and 155 of ADHD persons. figure 3 below is one of the Brain MRI scans of the actual ADHD individual.

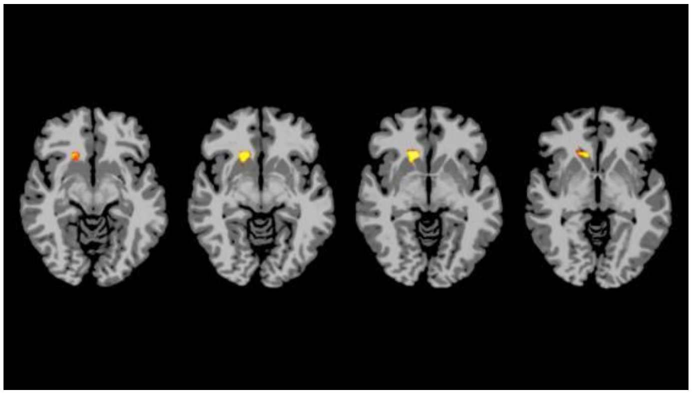

Figure 3: Brain scan with ADHD

3.3 Data Pre-processing

The data pre-processing techniques are as follows:

1. Missing Data: All the data which was missing in dataset was taken care of by using imputer class imported via Skcit.learn. NaN was filled in the place of missing data.

2. Encoding Categorical Data: All the data in the dataset is not always of same category to resolve this issue we use label encoding. In this we categorize the data into numeric form which can be applicable to model. Label encoder gives integer value to the data which is not in integer form.

3.Feature Scaling: Some of the values in dataset maybe in between 1-10 and some may be in 1000 so this will create problem and a biasness factor to higher value feature will be there. To tackle this feature scaling is used to normalize all the value in dataset using standardization or Normalization [26].

$$X' = \frac{X-\mu}{\sigma}$$

Standardization

$$X' = \frac{X-X_{min}}{X_{max}-X_{min}}$$

Normalization

a. Standardization: Here represents the mean of values in feature and means the feature value's standard deviation.

b. Normalization: Here Xmax is the max value in a feature whereas Xmin is min value.

In Deep learning pre-processing is as follows:

For the deep learning we have used VGG16 model and brain scans images dataset. There are all over 155 images belonging to ADHD class and 98 image belonging to non ADHD class.

First we divide the dataset in train set, test set and validation set. The augmentation started with image resizing, RGB images are converted to gray scale. Then the images are cropped so as to reduce the background noise which in turn will decrease the size of image pixel array. Data set so obtained is saved and passed to image generator. Image generator will automatically assign the classes to the dataset.

Existing approach models mostly all the researchers used SVM (kernel / simple), combinations of SVM with some other machine learning method. Very few researchers used deep learning models and that to complex such as 3D CNN and LSTM and maximum accuracy with SVM was 80%, with SVM RBF was 98% and 3D CNN was 76.6%

Proposed approach is more focused towards extracting best accuracy with the much simpler machine learning and deep learning models. It used ensemble machine learning models such bagging and boosting and CNN VGG-16 model that is much simpler and smaller in terms of memory requirement. Maximum accuracy of 89% is achieved by gradient oosting and 90% is achieved by VGG-16.

4. REULTS AND CONCLUSION

With the aim to contribute for the development of the society we opted to perform the research on the neural disorder namely Attention Deficit Hyperactivity disorder (ADHD). Several researches had been carried out to help the doctors detect and cure this neural disorder. We have applied different ensemble learning models like Extra tree classifier, Bagging and Ada boosting in order to improve the prediction and classification of disease and then at last we have applied the deep learning model CNN VGG-16 on Brain MRI scans to predict the ADHD disease and got the accuracy of 90%. In different researches different models are combined with SVM as one of the important parts to effectively produce the results for example CNN + SVM (linear/kernel).

Table 1 : CNN VGG16 Confusion Matrix

(0,'NO')	17	2
(1,'YES')	3	28
	(0,'NO')	(1,'YES')

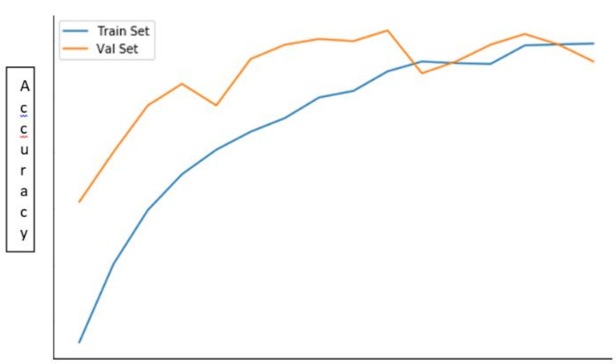

Figure 4 : CNN VGG-16 Model Accuracy

Table 2 : Accuracies of different classifiers and models

Model	Weighted average Precision	Weighted average Recall	Weighted average F1-Score	Accuracy
KNN	80%	80%	80%	80%
SVM	84%	84%	84%	84%
ETC	85%	85%	84%	85%
Bagging Classifier	83%	82%	83%	82%
Boosting Classifier	89%	89%	89%	89%
CNN VGG 16	90%	90%	90%	90%

Hence, we would like to establish that ensemble learning methods are more approachable process in the field of ADHD prediction as they secure good accuracies and takes less time and moderate software requirements and is very much low in costs. Though we get highest accuracy of 90% through CNN VGG 16 but it has taken more time and requires higher software requirements. Table 1 shows CNN VGG16confusion matrix. Figure 4 is the model accuracy of CNN VGG-16.

5. FUTURE SCOPE

In future, there is a scope to use a larger dataset so that this result is more reliable. Adapting additional parameters can give a scope of improvement of accuracy with a larger dataset including more factors which are affecting ADHD patients now needs to be inculcated in the dataset. It can be used as a foundation to additional research for similar disorder, A whole software which can be operated by doctors in hospitals [25] is what we aim for now.

REFERENCES

1. TJensen, P. S., Martin, D., & Cantwell, D. P., **Comorbidity in ADHD: implications for research, practice, and DSM-V** in Journal of the American Academy of Child & Adolescent Psychiatry, 36(8), pp. 1065-1079, 1997.
2. Singh, I., **Biology in Context: Social and Cultural Perspectives on ADHD. Research Review** in Children & society, 16(5), pp. 360-67, 2002 https://doi.org/10.1002/chi.746
3. Sinha, A., Pratik, R., & Panchal, M., **MRI IMAGE SEGMENTATION–A REVIEW**, 2014
4. Tian, L., Jiang, T., Wang, Y., Zang, Y., He, Y., Liang, M., ... & Zhuo, Y., **Altered resting-state functional connectivity patterns of anterior cingulate cortex in adolescents with attention deficit hyperactivity disorder** in Neuroscience letters, 400(1-2), pp. 39-43, 2006.
5. Simon, V., Czobor, P., Balint, S., Meszaros, A., Murai, Z., & Bitter, I., **Detailed review of epidemiologic studies on adult Attention Deficit/Hyperactivity Disorder (ADHD)** Psychiatria Hungarica: A Magyar Pszichiatriai Tarsasag tudomanyos folyoirata, 22(1), pp. 4-19, 2007.
6. Brown, M. R., Sidhu, G. S., Greiner, R., Asgarian, N., Bastani, M., Silverstone, P. H., ... & Dursun, S. M., **ADHD-200 Global Competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements** in Frontiers in systems neuroscience, 6, 69, 2012. https://doi.org/10.3389/fnsys.2012.00069

7. Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. J., **Functional connectivity: the principal - component analysis of large (PET) data sets** in Journal of Cerebral Blood Flow & Metabolism, 13(1), pp. 5-14, 1993.
8. Wang, S. H., Zhang, Y. D., Dong, Z., & Phillips, P., **Image Preprocessing for Pathological Brain Detection** in Pathological Brain Detection, Springer, Singapore, pp. 29-44, 2018.
9. Ray, S., **A Quick Review of Machine Learning Algorithms** in 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), pp. 35-39, IEEE.
10. Suykens, J. A., & Vandewalle, J. **Least squares support vector machine classifiers** in Neural processing letters, 9(3), pp. 293-300, 1999.
11. Paulus, M. P., Kuplicki, R., & Yeh, H. W., **Machine Learning and Brain Imaging: Opportunities and Challenges** in Trends in neurosciences, 42(10), pp. 659-661, 2019.
12. Badrinarayanan, V., Kendall, A., & Cipolla, R., **Segnet: A deep convolutional encoder-decoder architecture for image segmentation**in IEEE transactions on pattern analysis and machine intelligence, 39(12), pp. 2481-2495, 2017
13. Skurichina, M., & Duin, R. P., **Bagging, boosting and the random subspace method for linear classifiers** Pattern Analysis & Applications, 5(2), pp. 121-135, 2002.
14. Goetz, M., Weber, C., Bloecher, J., Stieltjes, B., Meinzer, H. P., & Maier-Hein, K., **Extremely randomized trees based brain tumor segmentation** in Proceeding of BRATS challenge-MICCAI, pp. 006-011, 2014.
15. Belaid, O. N., & Loudini, M., **Classification of Brain Tumor by Combination of Pre-Trained VGG16 CNN** in Journal of Information Technology Management, pp. 13-25, 2020.
16. Fcon_1000.projects.nitrc.org. 2020. ADHD200. [online] Available:<http://fcon_1000.projects.nitrc.org/indi/adhd200/> [Accessed 19 May 2020].
17. Kaur, S., Singh, S., Arun, P., Kaur, D., & Bajaj, M., **Phase Space Reconstruction of EEG Signals for Classification of ADHD and Control Adults**, in Clinical EEG and Neuroscience, 51(2), pp. 102-113, 2020.
18. Wang, Z., Sun, Y., Shen, Q., & Cao, L., **Dilated 3D Convolutional Neural Networks for Brain MRI Data Classification** IEEE Access, 7, pp. 134388-134398, 2019.
19. Wang, Tianyi, and Sei-ichiro Kamata. **Classification of Structural MRI Images in Adhd Using 3D Fractal Dimension Complexity Map** 2019 IEEE International Conference on Image Processing (ICIP). IEEE.
20. Tariq, S., Akhtar, N., Afzal, H., Khalid, S., Mufti, M. R., Hussain, S., ... & Ahmad, G., **A Novel Co-Training-Based Approach for the Classification of Mental Illnesses Using Social Media Posts**, IEEE Access, 7, pp. 166165-166172, 2019.
21. Alchalabi, A. E., Shirmohammadi, S., Eddin, A. N., & Elsharnouby, M., **FOCUS: detecting ADHD patients by an EEG-based serious game**, in IEEE Transactions on Instrumentation and Measurement, 67(7), pp. 1512-1520, 2018.
22. Wong, D. F., Ng, T. K., Ip, P. S., Chung, M. L., & Choi, J., **Evaluating the Effectiveness of a Group CBT for Parents of ADHD Children. Journal of Child and Family Studies**, 27(1), pp. 227-239, 2018.
23. Gamma, Alex, and Olga Kara, **Event-Related Potentials for Diagnosing Children and Adults with ADHD.**In Journal of attention disorders, 1087054716631821, 2016.
24. Cuffe, S. P., Visser, S. N., Holbrook, J. R., Danielson, M. L., Geryk, L. L., Wolraich, M. L., & McKeown, R. E., **ADHD and psychiatric comorbidity: Functional outcomes in a school-based sample of children** in Journal of attention disorders, 1087054715613437, 2015.
25. Munya A. Arasi , Sangita babu, **Survey of Machine Learning Techniques in Medical Imaging,** International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.5, ISSN 2278-3091, September - October 2019. https://doi.org/10.30534/ijatcse/2019/39852019
26. Randhir Singh , Saurabh Pal, **Machine Learning Algorithms and Ensemble Technique to Improve Prediction of Students Performance,** International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.5, ISSN 2278-3091, May - June 2020. https://doi.org/10.30534/ijatcse/2020/221932020
27. Saeed, Fahad. "Towards quantifying psychiatric diagnosis using machine learning algorithms and big fMRI data." in Big Data Analytics 3.1, 7, 2018.
28. Yao, D., Guo, X., Zhao, Q., Liu, L., Cao, Q., Wang, Y., ... & Sui, J., "**Discriminating ADHD from healthy controls using a novel feature selection method based on relative importance and ensemble learning**" in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4632-4635, IEEE.
29. Brown, M. R., Sidhu, G. S., Greiner, R., Asgarian, N., Bastani, M., Silverstone, P. H.,& Dursun, S. M., "**ADHD-200 Global Competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements**" in Frontiers in systems neuroscience, 6, 69, 2012.
30. Zou, Liang, Jiannan Zheng, and Martin J. McKeown, "**Deep learning based automatic diagnoses of attention deficit hyperactive disorder**" in 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP),IEEE.