**Indian Institute of Information Technology, Allahabad**

**Semester VI**

DETECTION OF MAJOR DEPRESSIVE DISORDER(MDD) AND ATTENTION DEFICIT HYPERACTIVITY DISORDER(ADHD) IN CHILDREN, TEENS AND YOUNG ADULTS

**Project Report**



**Group 5:**

Pushkal Madaan IIT2020005

Ritej Dhamala IIT2020006

Tanushree IIT2020054

Anushka Dandawate IIT2020077

Rushil Patra IIB2020005

## **Table of Contents**

## 

| **Sl.** | **Topic** | **Page** |
| --- | --- | --- |
| 1 | Abstract | 2 |
| 2 | Keywords | 2 |
| 3 | Introduction | 2 |
| 4 | Literature review | 3 |
| 5 | Dataset description | 5 |
| 6 | Methodology | 6 |
| 7 | Result and Discussion | 11 |
| 8 | Conclusion and Future Work | 13 |
| 9 | References | 14 |

**Abstract**

In this project, our aim is to develop a classification model to identify Major Depressive Disorder (MDD) and Attention Deficit Hyperactivity Disorder (ADHD) in young teens and adults, which often go undetected. We have trained and tested several classifiers, such as K-Nearest Neighbors (KNN), Logistic Regression, Support Vector Machine (SVM), and a hybrid ensemble model of SVM, KNN, and Logistic Regression on the TD EEG Brain dataset. We have compared the results with various other papers to evaluate the effectiveness of our model.

**Keywords**

Attention Deficit Hyperactivity Disorder (ADHD), Major Depressive Disorder (MDD), electroencephalogram (EEG)

**Introduction**

In India, students are under immense academic pressure, leading to a rise in cases of ADHD and depression, the two most common mental illnesses. The recent COVID pandemic has made it worse, causing early depression in many children. ADHD prevalence is high in children aged 8 to 15, with a range of 9.40% in male children and 5.20% in female children. However, many go undiagnosed, and leaving ADHD untreated can lead to Dysthymia.[1] Clinical assessments and interviews conducted by experts like psychiatrists and psychologists are used to diagnose depression, but a different strategy is required due to varying symptoms and children's inability to articulate their feelings.

Electroencephalography (EEG) data can aid in the diagnosis of these disorders as it provides valuable insights into the functioning of the brain.[2] In this report we have made use the TD brain dataset which contains EEG recordings of healthy individuals and individuals with ADHD and MDD, and have machine learning models to predict the presence of these disorders. Early detection and treatment of MDD/ADHD can increase academic and occupational performance, lower the risk of substance dependence, improve mental health outcomes, and improve social ties and general life quality.

In this study, we investigate the application of machine learning methods for the identification of MDD and ADHD in children, adolescents, and young adults. Based on brain eeg data, we examine the effectiveness of a number of machine learning models, such as support vector machines (SVMs), K Nearest Neighbour, SVM, and a novel approach of combining these models using hybrid ensemble learning. Our findings show that machine learning algorithms can offer precise and trustworthy diagnostic assessments for these conditions, and they may even be able to enhance patient outcomes in terms of early detection and treatment.

**Related work**

The classification and identification of Major Depressive Disorder (MDD) and Attention Deficit Hyperactivity Disorder (ADHD) in young adolescents and adults is a critical area of research in the field of mental health.[3] Various machine learning techniques, including feature extraction utilizing statistical methods and classifiers like KNN, logistic regression, SVM, and hybrid ensemble models, were used to do this.

[4]A paper published in IEEE studies the use of machine learning algorithms to predict ADHD in children using a dataset of 157 children, out of which 77 had ADHD and 80 were healthy. The study evaluates the performance of three algorithms - Naïve Bayes, kNN, and Logistic Regression. The kNN algorithm performs the best, with an accuracy of 86%, followed by logistic regression at 66%, and Naïve Bayes at 52%. The kNN algorithm is able to predict ADHD with an accuracy of 89%. The study highlights the potential of machine learning in predicting ADHD and identifies kNN as the best performing algorithm for this task.

[5]Another study published in the IEEE presents a machine learning-based approach for the early detection and classification of children with ADHD using EEG signals. The study uses two separate approaches to identify optimal channels and important features, and a hybrid channel selection method to combine these approaches. Six machine learning classifiers are then applied to detect ADHD in 121 children aged 7-12 years, using only six channels and twenty-eight features. The Gaussian process-based classifier achieved an accuracy rate of 97.53% and an AUC of 0.999, an improvement over previously developed techniques. The proposed system can be helpful for doctors and physicians to detect children with ADHD early and provide appropriate healthcare services and treatment.

[6]Another study published in IEEE focused on developing an end to end deep model.The goal of this work was to create a deep learning (DL) model for categorising people with Major Depressive Disorder (MDD) and healthy controls using resting-state electroencephalography (EEG) data. The suggested model learned connection links among EEG channels using a multi-head self-attention mechanism, extracted higher-level features using a parallel two-branch convolutional neural network, and classified using a fully connected layer. On a publicly accessible EEG dataset, the model outperformed comparative approaches, with an average classification accuracy of 91.06%. These findings show that the proposed DL model provides a feasible technique for MDD identification using brain connection modelling.

[7]Some researchers proposed to identify major depressive disorder by selecting discriminative features via stochastic search.The goal of this study was to create an objective biomarker-based technique for diagnosing Major Depressive Disorder (MDD) utilizing electroencephalography (EEG) data. To identify discriminative characteristics from specific EEG channels, the suggested method used a stochastic search algorithm. The approach was tested using a public EEG-based MDD dataset that included 24 MDD patients and 29 healthy controls. The protocol of leave-one-subject-out cross-validation was used. The suggested method outperformed state-of-the-art MDD recognition systems, with an average accuracy of 99.53% in the fear-neutral face pairs experiment and 99.32% in the resting condition. The study also discovered that negative emotional inputs can cause depressive moods and that high-frequency EEG patterns can distinguish between normal and depressive patients.

In our study, we used statistical feature extraction methods and classifiers such as KNN, logistic regression, SVM, and a hybrid ensemble of KNN, logistic regression, and SVM to classify MDD and ADHD. The statistical feature such as mean, standard deviation, min, max, skewness, etc. and hybrid ensemble learning to classify eeg data is the novel approach we are proposing though this paper. We believe combining the best tunned parameter of stand alone learning methods into a hybrid ensemble leanring will effectively enchance the classification accuracy of the eeg data, while statistical feature wil help the model to learn hidden characteristics of our data. Our findings are compared to those of other studies, such as Patel et al. and Shi et al.

Overall, the studies reviewed in this literature review demonstrate that machine learning techniques can effectively classify and identify MDD and ADHD using various neuroimaging and EEG data. These findings highlight the potential of machine learning techniques to aid in early detection and accurate diagnosis of mental health disorders, leading to better outcomes for individuals and society as a whole.

**Dataset description**

The TD EEG Brain Dataset is a significant resource for academics interested in investigating brain function and neuroimaging in normal development[8]. It is made up of EEG data taken from typically developing children and young adults, and it contains data from 12 different brain regions recorded during rest and task circumstances such as visual and auditory stimuli, cognitive tasks, and motor tasks.

The collection comprises EEG data from approximately 150 subjects ranging in age from 6 to 18 years, as well as complete demographic information, cognitive assessment scores, and behavioral measures for each participant. This enables researchers to explore the links between brain activity, behavior, and cognition, as well as get a better knowledge of typical brain development.

The availability of this dataset has allowed collaborative research efforts and has been used in a variety of brain development research investigations. These studies involve research into neural oscillations, functional connectivity, and the effects of cognitive and sensory processing on brain activity.

The use of the TD EEG Brain Dataset in research has yielded vital insights into the development of brain function in typically developing children and young people. For example, studies using this dataset have revealed differences in brain activity between children and adults during cognitive activities and neural markers of attention and cognitive processing. These findings have crucial implications for understanding brain development and devising therapies to improve cognitive growth in children and young people.

**Methodology**

**Introduction**

The data contained recordings from 26 electrodes or channels for each patient. To optimize our prediction model, we employed two different approaches. The first approach used the recordings from all 26 electrodes, while the second approach involved selecting only the relevant electrode recordings.

**Electrode Filtering**

The frontal lobes of the brain are responsible for many important cognitive processes, including attention, memory, decision-making, and emotion regulation. These processes are often disrupted in individuals with ADHD and MDD. Research such as [15] has shown that individuals with ADHD have decreased activity in the prefrontal cortex, a region located in the frontal lobes, which is associated with attention, impulsivity, and executive function. Similarly, individuals with MDD often show decreased activity in the prefrontal cortex, which is involved in emotion regulation and cognitive control. Both ADHD and MDD have been associated with dysfunction in the prefrontal cortex, which is a part of the frontal lobe.

Therefore, EEG recordings from electrodes placed on the frontal lobe can provide important information on the neural activity in this region and can potentially aid in the diagnosis of these disorders. Additionally, EEG recordings from the prefrontal cortex can be used to monitor the effects of treatment on brain activity in individuals with ADHD or MDD.

Hence the electrodes we filtered out are -'Fp1', 'Fp2', 'F3', 'F4', 'F7', 'F8', 'Fz'

**Data preprocessing**

The data in CSV files is imported and processed using Pandas, and then transformed into a MNE RawArray object to enable further processing. As data collected over time in neuroimaging and electrophysiology is difficult to analyze as a continuous signal, it is segmented into short, contiguous segments of data called epochs.

**Epochs and Overlaps**

Epochs are typically centered around specific events or stimuli and help in breaking down the continuous EEG data into smaller, more manageable segments. This allows for easier identification of patterns or anomalies, and the use of fixed-length epochs ensures that data is consistent across different subjects and recording sessions, facilitating the comparison of data. By dividing the continuous signal into epochs, it becomes simpler to perform analysis on specific events or features of the data.

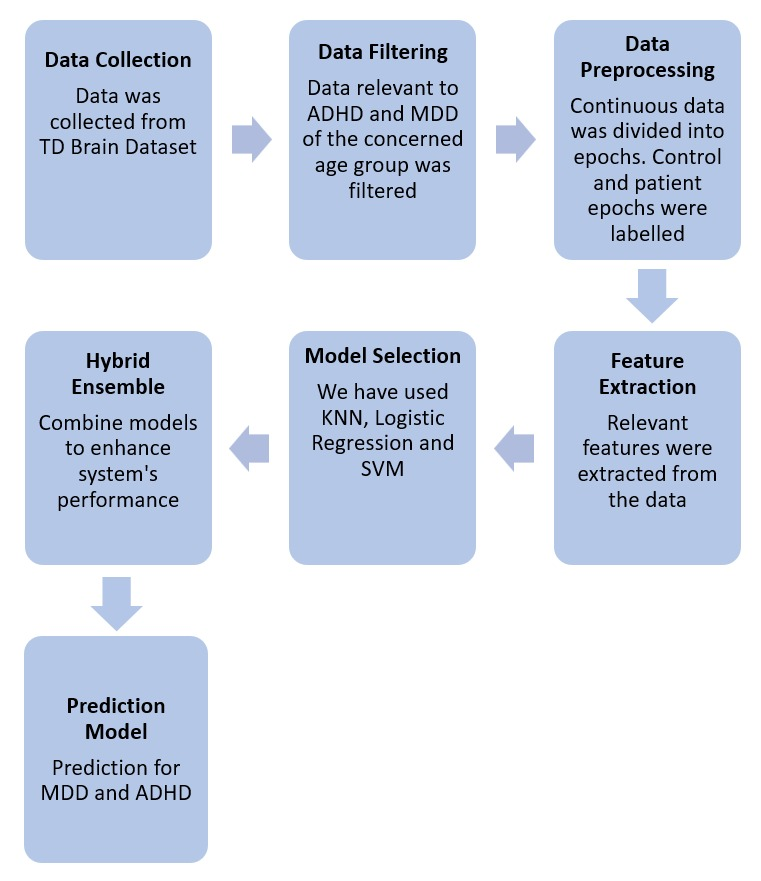
Overlapping the epochs decreases the effect of edge artifacts. These artifacts may arise from sudden changes in the signal at the beginning or end of an epoch. Overlapping the epochs can decrease the impact of these artifacts, which may enhance the accuracy of our analysis. Furthermore, overlapping epochs can expand the dataset size without requiring additional recordings. The use of overlaps has numerous benefits, such as minimizing data loss and increasing the number of samples. By incorporating data from the prior epoch into the current epoch, overlaps can help prevent important information from being lost. Additionally, more data can enhance the accuracy of machine learning models.

It is crucial to choose the appropriate epoch duration and overlap since it impacts the quality of the features extracted and, consequently, the performance of the classification algorithm. A duration that is too short may not capture enough information about the underlying brain activity, while a duration that is too long may include various neural processes and result in a loss of specificity. Similarly, an overlap that is too small may cause the epochs to miss critical information, while an overlap that is too large may generate data points that are too alike and do not provide enough independent information for accurate classification.

**Feature Extraction**

We have used a plethora of time-series and morphologically-based features for our models. These features are extracted by the analysis of the shape/form of the signals where we perform a transformation on a particular signal and extract relevant measurements on that transformed signal which will perform as our features. The features extracted from the TDBrain Dataset include:

* Mean
* Standard Deviation
* Peak to Peak (pk-pk)
* Variance
* Minimum Amplitude
* Maximum Amplitude
* Minimum Amplitude for a given argument (Argmin)
* Maximum Amplitude for a given argument (Argmax)
* Absolute difference between signals
* Skewness
* Kurtosis



**KNN**

The KNN technique places an input signal in the category that is most common among its K nearest neighbours after finding the input signal's K nearest neighbours.The retrieved feature vectors have been subjected to the KNN algorithm. To get the highest classification performance, we have tweaked the hyperparameters K, distance measure, and regularisation. Better generalisation can be achieved with a higher value of K, but the algorithm's computing cost may also rise.

**Logistic Regression**

The logistic regression approach uses a sigmoid function to represent the connection between the collected characteristics and the binary output variable (class label). The projected probabilities are thresholded at 0.5, which indicates the likelihood that the EEG signal belongs to the healthy class (0) or the patient class (1). This allows the sigmoid function to transfer the output of the linear function to a value between 0 and 1.

**SVM**

SVM produces a hyperplane to divide the feature vectors by maximising the distance between the hyperplane and the support vectors, or nearest neighbour feature vectors, of each class. The margin is the separation between the closest support vectors from each class and the hyperplane. By minimising the margin violation, which is the total of the distances between the incorrectly classified data points and the hyperplane, the SVM algorithm discovers the ideal hyperplane. SVMs can use kernel functions to change the feature space into a higher-dimensional space where the data can be linearly separable in situations where the data is not linearly separable. In order to select the top kernel function from a variety of kernel functions to be employed in our classification, we used the gridsearch approach.

**Hybrid Ensemble Learning**Hybrid ensemble learning is a method that mixes various machine learning models to enhance the system's overall performance. Using this method, the same dataset is used to train numerous classifiers, and the results are pooled to get the final prediction. Voting, weighting, or stacking are a few methods that can be used to accomplish this. We used the same dataset to train SVM, KNN, and logistic regression classifiers, and we combined the results via stacking and voting.

**Voting Classifier**

Voting classifier combines the final prediction with the predictions of various independent classifiers.Hard voting and soft voting are two examples of the various voting classifications. In a hard vote, the outcome is determined by the classifiers who received the most votes overall. The final prediction in soft voting is based on the mean of the anticipated probabilities from each classifier. After testing both approaches, we chose hard voting since it consistently outperformed the others in all of our tests.

**Stacking Classifier**

Stacking classifier is another hybrid ensemble learning method that integrates the results of various separate classifiers using a meta-classifier. By training a number of base classifiers on the same dataset using several feature sets and hyperparameters, it is possible to aggregate the predictions of various algorithms and feature sets. A meta-classifier is trained using the predictions of these base classifiers as input features which makes the final prediction. The meta-classifier can enhance the overall performance of the classification system by capturing the intricate connections between the base classifiers.

**Evaluation Criteria**

We have chosen four different evaluation methods to see how the different models hold up against each other as well as against models from other published studies. These comprise of:

Precision - Represents how many positive identifications were made from all the predicted positives.

Recall - Represents how many true positives were identified from all the positive identifications.

Accuracy - It is the measure of all correctly identified cases

F1 Score - This is the harmonic mean of recall and precision. It results in a stronger measure of the incorrectly classified cases compared to the accuracy metric.

**Results**

Among all the models' the Hybrid ensemble model performed the best in all experiments. Incorporating all 26 channels from all regions of the head instead of taking only 7, the frontal lobe, showed significant improvement in the detection of both diseases. Readings taken with eyes open were found to give slightly more accuracy than eyes closed ones. Considering single models KNN and SVM had comparable results. KNN performed better than SVM in ADHD 7-channel experiment with +1.6% accuracy. In another case, SVM turned better with slight margins. Logistic regression was nowhere near the other two models.

Below is the full comparison of all the models applied to experiment-specific filtered data.

| Models | Experiments | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Eyes Closed | | | | Eyes open | | | |
| MDD CH 7 | MDD CH 26 | ADHD CH 7 | ADHD CH 26 | MDD CH 7 | MDD CH 26 | ADHD CH 7 | ADHD CH 26 |
| KNN | 92.7 | 95.6 | 90.4 | 96.8 | 93.7 | 97.1 | 93.1 | 96.3 |
| Logistic Regression | 77.2 | 91.1 | 77.5 | 86.2 | 80.1 | 94.0 | 72.9 | 87.6 |
| SVM | 92.7 | 94.9 | 89.0 | 95.9 | 94.0 | 98.1 | 95.9 | 96.3 |
| Voting Classifier | 91.8 | 95.9 | 90.4 | 95.0 | 93.0 | 97.6 | 94.5 | 97.2 |
| Stacking Classifier | 94.0 | 96.2 | 94.0 | 96.8 | 94.9 | 98.3 | 95.0 | 96.8 |

*Tabel 1: Comparison of model with different data*

| Paper | Models used | Accuracy |
| --- | --- | --- |
| doi:10.1109/ICACCS54159.2022.9785356 [5] | K Nearest Neighbor | 86% |
| Naïve Bayes | 52% |
| logistic regression | 66 % |
| doi:10.1109/ACCESS.2023.3264266 [6] | Gaussian process classification (GPC) | 97.53% |
| Random forest | 75.17% |
| Multilayer perceptron | 95.03% |
| doi:10.1109/ACCESS.2023.3270426 [7] | CNN-LSTM | 84.9% |
| DeprNet | 87.7% |
| Proposed mode | 90.5% |
| doi:10.1088/1741-2552/acbe20 [8] | CNN | 95.49% |
| DCNN-LSTM | 99.24% |
| RSSA-BLDA | 99.53% |
| **Our Proposed Model (MDD) (EC)** | **Hybrid Ensemble Voting Classifier (Logistic Regression, KNN, SVM)** | **98.1%** |
| **Hybrid Ensemble Stacking Classifier (Logistic Regression, KNN, SVM)** | **97.6%** |
| **Our Proposed Model (ADHD) (EC)** | **Hybrid Ensemble Voting Classifier (Logistic Regression, KNN, SVM)** | **96.3%** |
| **Hybrid Ensemble Stacking Classifier (Logistic Regression, KNN, SVM)** | **97.2%** |

*Table 2: Comparison with related papers*

**Conclusion and future work**

In conclusion, detecting Major Depressive Disorder (MDD) and Attention Deficit Hyperactivity Disorder (ADHD) in children, teens, and young adults is crucial for effective treatment and management of these conditions. The symptoms of MDD and ADHD can vary widely between individuals, but some common signs include changes in mood, behavior, and attention span.

Early detection is important to prevent long-term negative consequences, such as poor academic and social performance, substance abuse, and increased risk of suicide. A comprehensive evaluation by a qualified mental health professional is essential to accurately diagnose MDD and ADHD, as well as rule out other possible conditions that may present with similar symptoms.

It's important to note that developing an effective classification model for mental health disorders such as MDD and ADHD requires a large and diverse dataset, rigorous testing, and validation. Therefore, it's essential to ensure that the model's accuracy and effectiveness are rigorously evaluated before it is implemented in real-world clinical settings.

We have observed other existing eeg classification mechanisms like doi:10.1088/1741-2552/acbe20[8] has slightly better performance than our model. Upon inspection we have found that the non linear feature used by the authors is the reason their classifier was able to perform better. This means if we can extract and incorporate non linear feature into our model, the performance of our model can be improved which we will try to include in our future work.

Future work should focus on addressing the systemic barriers and stigmas that prevent individuals from seeking help and receiving proper treatment. This can involve improving access to mental health services, reducing the cost of treatment, and increasing awareness and education about MDD and ADHD in schools, workplaces, and communities.

Overall, continuing the work on the detection and treatment of MDD and ADHD in children, teens, and young adults is critical for improving the mental health outcomes and quality of life for individuals affected by these conditions.

**References:**

[1] S. S. Poorna et al., "Classification of EEG based control using ANN and KNN — A comparison," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, India, 2016, pp. 1-6, doi: 10.1109/ICCIC.2016.7919524

[2] Movahed, Reza Akbari et al. “A major depressive disorder classification framework based on EEG signals using statistical, spectral, wavelet, functional connectivity, and nonlinear analysis.” *Journal of neuroscience methods* vol. 358 (2021): 109209. doi:10.1016/j.jneumeth.2021.109209

[3] G. Ciodaro, A. J. Najafabadi and B. Godde, "Resting State EEG Classification of Children With ADHD," 2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, USA, 2020, pp. 1-6, doi: 10.1109/SPMB50085.2020.9353628.

[4] S. Saini, R. Rani and N. Kalra, "Prediction of Attention Deficit Hyperactivity Disorder (ADHD) using machine learning Techniques based on classification of EEG signal," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2022, pp. 782-786, doi: 10.1109/ICACCS54159.2022.9785356.

[5] M. Maniruzzaman, M. A. M. Hasan, N. Asai and J. Shin, "Optimal Channels and Features Selection Based ADHD Detection From EEG Signal Using Statistical and Machine Learning Techniques," in IEEE Access, vol. 11, pp. 33570-33583, 2023, doi: 10.1109/ACCESS.2023.3264266.

[6] M. Xia, Y. Zhang, Y. Wu and X. Wang, "An End-to-End Deep Learning Model for EEG-Based Major Depressive Disorder Classification," in IEEE Access, vol. 11, pp. 41337-41347, 2023, doi: 10.1109/ACCESS.2023.3270426.

[7] Chang H, Zong Y, Zheng W, Xiao Y, Wang X, Zhu J, Shi M, Lu C, Yang H. EEG-based major depressive disorder recognition by selecting discriminative features via stochastic search. J Neural Eng. 2023 Mar 23;20(2). doi: 10.1088/1741-2552/acbe20. PMID: 36812637.

[8] van Dijk, H., van Wingen, G., Denys, D. et al. The two decades brainclinics research archive for insights in neurophysiology (TDBRAIN) database. Sci Data 9, 333 (2022). https://doi.org/10.1038/s41597-022-01409-z

[9] Mayer, J.S., Brandt, G.A., Medda, J. *et al.* Depressive symptoms in youth with ADHD: the role of impairments in cognitive emotion regulation. *Eur Arch Psychiatry Clin Neurosci* 272, 793–806 (2022).

[10] Michael C. Meinzer, Jeremy W. Pettit, Chockalingam Viswesvaran, The co-occurrence of attention-deficit/hyperactivity disorder and unipolar depression in children and adolescents: A meta-analytic review, Clinical Psychology Review, Volume 34, Issue 8, 2014.

[11] Diler, Rasim Somer et al. “Differentiating major depressive disorder in youths with attention deficit hyperactivity disorder.” *Journal of affective disorders* vol. 102,1-3 (2007): 125-30. doi:10.1016/j.jad.2007.01.002

[12] Sadeghian Nadooshan, M., Shahrivar, Z., Mahmoudi Gharaie, J. *et al.* ADHD in adults with major depressive or bipolar disorder: does it affect clinical features, comorbidity, quality of life, and global functioning?. *BMC Psychiatry* 22, 707 (2022).

[13] McIntyre, Roger S., et al. "Attention-deficit/hyperactivity Disorder in Adults With Bipolar Disorder or Major Depressive Disorder: Results From the International Mood Disorders Collaborative Project." *Primary Care Companion to the Journal of Clinical Psychiatry,* vol. 12, no. 3, 2010.

[15] Kibby MY, Dyer SM, Lee SE, Stacy M. Frontal volume as a potential source of the comorbidity between attention-deficit/hyperactivity disorder and reading disorders. Behav Brain Res. 2020 Mar 2;381:112382. doi: 10.1016/j.bbr.2019.112382. Epub 2020 Jan 7. PMID: 31917238; PMCID: PMC7064054.