

Capstone Project Phase A

**Applying graph theory measures for analyzing networks in brain recordings.**

**PROJECT CODE:**

**24-1-R-6**

**Supervisor:**

**Dr. Anat Dahan**

**Authors:**

**Koral Pnina Biton, ID: 314860057**

[coralbiton20@gmail.com](mailto:coralbiton20@gmail.com)

**Topaz Eldori, ID: 319003141**

[eldoritopaz85@gmail.com](mailto:eldoritopaz85@gmail.com)**‬**

*Table of Contents*

[1. Abstract. 3](#_Toc163329879)

[2. Introduction 4](#_Toc163329880)

[3. Background and related work 5](#_Toc163329881)

[*3.1 ADHD Recognition 5*](#_Toc163329882)

[*3.2 EEG Brain Recording 5*](#_Toc163329883)

[*3.3 The Relationship Between Graph Theory and Brain Recordings 7*](#_Toc163329884)

[*3.4 Community Detection 11*](#_Toc163329885)

[4. Expected Achievements 12](#_Toc163329886)

[5. Engineering Process 13](#_Toc163329887)

[*5.1 Process 13*](#_Toc163329888)

[*5.2 Product 14*](#_Toc163329889)

[***5.2.1 Use Case Diagram 14***](#_Toc163329890)

[***5.2.2 Data Process*** ***15***](#_Toc163329891)

[***5.2.3 Prototype 21***](#_Toc163329892)

[6. Verification plan 21](#_Toc163329893)

[7. Reference: 21](#_Toc163329894)

# ***Abstract.***

*This project explores the application of graph theory measures to analyze brain networks using EEG recordings. It investigates differences in information pathways within brain networks between children with attention deficit hyperactivity disorder (ADHD) and healthy subjects. EEG recordings from both groups during an attentional visual task were analyzed to calculate effective connectivity. Group differences in connectivity between brain regions were evaluated, revealing disrupted patterns of information flow in ADHD children compared to healthy subjects. Specifically, differences were observed in posterior to anterior information flow in the theta band and in pathways between anterior regions in the beta band. The study sheds light on the altered information flow patterns in ADHD children compared to their healthy counterparts, providing potential insights into the neural mechanisms underlying ADHD.*

**Keywords***: Hyperactivity Disorder (ADHD), Electroencephalography (EEG), Brain networks, Graph theory, Functional connectivity, Information pathways, Neurodevelopmental disorders, Computational neuroscience, Therapeutic targets, Community detection, Network analysis, Information processing, Cognitive neuroscience, Neural correlates.*

# **Introduction**

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder characterized by symptoms of inattention, hyperactivity, and impulsivity, often persisting into adulthood. With a prevalence rate of approximately 5% among children worldwide, ADHD represents one of the most common psychiatric disorders diagnosed in childhood. The disorder can have profound implications for academic performance, social interactions, and daily functioning, underscoring the importance of early diagnosis and intervention. ‎[13]

Electroencephalography (EEG) has emerged as a valuable tool in the assessment of brain function, offering insights into neural activity through the measurement of electrical signals generated by neurons. EEG recordings provide a non-invasive and objective means of evaluating brain activity, facilitating the identification of aberrant patterns associated with neurological disorders such as ADHD. By capturing the temporal dynamics of neuronal oscillations, EEG enables researchers to investigate the functional connectivity and information processing within brain networks. Graph theory offers a powerful framework for the analysis of complex networks, including those derived from EEG recordings. In the context of brain networks, nodes represent distinct brain regions, while edges denote functional or structural connections between these regions. Graph theory measures, such as clustering coefficient, global efficiency, and modularity, provide quantitative descriptors of network topology, illuminating the organization and dynamics of brain networks.

The integration of EEG recordings and graph theory analysis holds promise for elucidating the neural correlates of ADHD. By examining differences in brain network architecture between individuals with ADHD and neurotypical controls, researchers can gain valuable insights into the underlying pathophysiology of the disorder. Specifically, alterations in information processing, connectivity patterns, and network dynamics may be identified, providing a basis for understanding the cognitive and behavioral manifestations of ADHD.

This project endeavors to contribute to the burgeoning field of computational neuroscience by applying graph theory measures to analyze brain networks in individuals with ADHD. Through the investigation of EEG recordings obtained during attentional tasks, the study aims to delineate distinct information pathways within brain networks and elucidate the disruptions associated with ADHD. During our investigations we provide a comprehensive overview of the background literature pertaining to ADHD recognition, EEG brain recording techniques, the relationship between graph theory and brain networks, and existing methods for community detection in network analysis. We also outline the expected achievements of this research project and delineate the engineering process involved in data acquisition, preprocessing, and analysis.

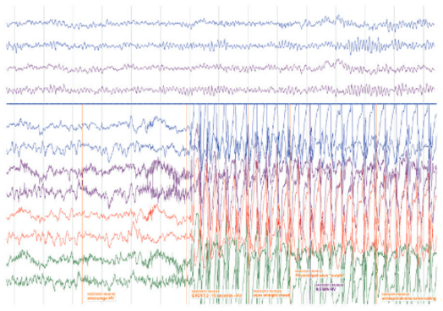
# **Background and related work**

## **3.1 ADHD Recognition**

Attention deficit hyperactivity disorder (ADHD) is one of the most common disorders, especially among children. The disorder is characterized by symptoms such as hyperactivity, difficulty paying attention, and impulsive behavior. Studies on the subject claim that about 5% of children have ADHD, with studies finding that boys are more affected. Symptoms can differ between children and typically appear in preschool age, increasing during school age. The symptoms can impact academic performance, social interactions, and daily activities. With earlier diagnosis of ADHD, there is reduced impairment in functioning. [‎13]

## **3.2 EEG Brain Recording**

An electroencephalogram (EEG) is a diagnostic technique used to measure brain activity by recording electrical signals generated by neurons. During an EEG procedure, small flat metal discs called electrodes are placed on the scalp to detect and record these electrical signals. The EEG recording displays a line for each electrode, showing how the activity detected by that electrode changes over time as it moves up and down. These lines can be analyzed to understand brain activity. EEG records only the brain's electrical activity without causing any discomfort or stimulation. It helps detect brain wave abnormalities, aiding in early diagnosis, treatment planning, and therapy monitoring.‎[14]



***Fig. 1: Example of EEG recording.*** *‎[14]*

The recorded waveforms reflect the cortical electrical activity.  
Signal frequency: the main frequencies of the human EEG waves are as follows:

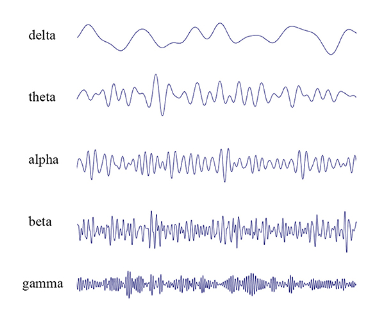
**Delta** **waves**: Delta waves occurring when you are during deep sleep, these slow brain waves are linked to the body's healing and restoration. These waves typically have a frequency of 1-4 Hertz**.** [‎16]

**Theta waves**: Theta waves occurring when you are during sleep or daydreaming when the mind is immersed in thought, are linked to relaxation and diminished alertness. These waves typically have a frequency of 4-8 Hertz. ‎[16]

**Alpha** **waves**: Alpha waves occurring when you are awake but in a relaxed state of mind, are linked to being calm and alert. They can enhance activities that require learning and coordination. These waves typically a frequency between 8 and 12 Hertz. ‎[16]

**Beta waves**: Beta waves occur throughout your daily activities, when you're awake, alert, busy, and focused. These waves typically a frequency between 12 to 38 Hertz and can be categorized into the following subtypes:

* **Low beta waves:** These waves typically have a frequency between 12 and 15 Hertz, they manifest during periods of active thinking.
* **Beta waves:** These waves typically have a frequency between 15 to 22 Hertz, these waves are prominent when you are concentrating on activities or tasks.
* **High beta waves:** These waves typically have a frequency between 22 to 38 Hertz, they arise during moments of excitement, anxiety, exposure to new experiences, or when dealing with complex thoughts. ‎[16]

**Gamma:** Gamma waves occur when you are super alert and wide awake. They have a frequency of 30 Hertz or higher, making them faster than any other brain waves. These waves are associated with intense focus and deep thinking. The effects of gamma waves depend on their level in your brain. If you have a lot of them, you're likely to feel happier and more open-minded. You might also find yourself more focused. However, if your brain lacks gamma waves, you might experience difficulties with learning and memory. In severe cases, a deficiency in gamma waves could lead to learning difficulties or mental health issues. **‎**[16**]**

***Fig. 2: EEG Waveforms.*** *[‎16]*

## **3.3 The Relationship Between Graph Theory and Brain Recordings**

Graph Theory in Network Representation:

Graph theory provides the mathematical framework to analyze networks and their information flow. Graph theory offers a framework for representing and understanding networks, ranging from social interactions to biological systems such as the brain. In this framework, networks are conceptualized as collections of nodes, which can represent various entities such as neurons, brain regions, or individuals. These nodes are interconnected by edges, symbolized by lines, which denote the relationships or connections between the nodes.

One essential characteristic of graph theory is its ability to capture different types of interactions within networks. For instance, edges in a graph can be directed, indicating the presence of a one-way relationship between nodes. This concept is particularly relevant in scenarios where information or influence flows asymmetrically, such as in directed social networks or neuronal signaling pathways.

Furthermore, graphs can also incorporate the notion of edge weights, which represent the strength or intensity of the connections between nodes. Weighted graphs are instrumental in capturing the relationships present in real-world networks. For instance, in a social network, the weight of an edge could signify the frequency of interaction between individuals, while in a neural network, it could represent the strength of synaptic connections between neurons. ‎[15]

**Measures In Graph Theory:**

In graph theory, both local and global measures are used for analysis. Local measures focus on individual nodes or edges, while global measures describe properties of the entire network.

Clustering Coefficient (Local): The local clustering coefficient of a node measures how close its neighbors are to being a complete graph. **‎**[16**]**

Strength (Local): This is the sum of the weights of all edges connected to a node, indicating the importance of a node in terms of weighted connections. **‎**[16**]**

Average Degree (Global): This is the average number of connections each node has in the network. **‎**[16**]**

Characteristic Path Length (Global): The average distance between any two nodes in the network. It's calculated by finding the average of the shortest paths between all pairs of nodes. **‎**[16**]**

Clustering Coefficient (Global): The average of the local clustering coefficients of all nodes, indicating how clustered the network. **‎**[16**]**

Global Efficiency (Global): This measure how efficiently information can be transmitted across the entire network. It's the average of the shortest path lengths between all pairs of nodes. **‎**[16**]**

Small-Worldness (Global): A small-world graph has a similar average path length to a random graph with the same degree distribution but is significantly more clustered. **‎**[16**]**

Transitivity (Global): This is the ratio of the total number of triangles (sets of three nodes where each is connected to the other two) to the number of all possible triplets of nodes in the network. **‎**[16**]**

Assortativity Coefficient (Global): It indicates whether nodes with similar degrees tend to be connected to each other. A positive AC means nodes with similar degrees are more likely to be connected. **‎**[16**]**

Modularity (Global): This measure how well the network can be divided into distinct communities or modules. It's determined based on a pre-defined community structure and reflects the quality of such division. **‎**[16**]**

***High modularity***

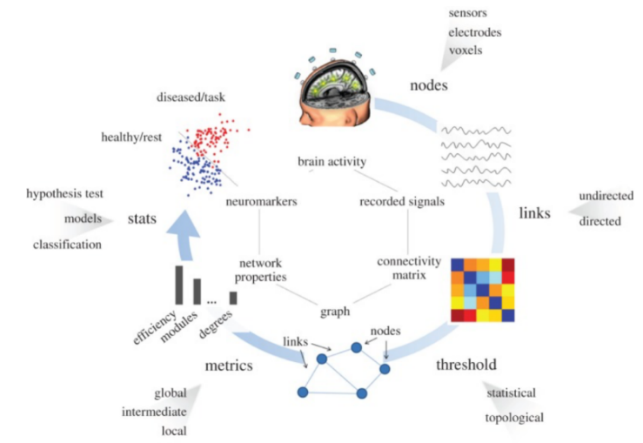
***Low modularity***



***Fig. 3: Modularity Classification.*****‎**[16**]**

Graph theory and brain recording:

Our brain functions like a network, with different parts working together during tasks and neurological states. Electroencephalography (EEG) testing is important for understanding brain activity. By using graph theory, we can analyze EEG data to reveal the complex communication networks within the brain. In this approach, nodes represent the electrode locations, while edges indicate functional connectivity between these locations. With this method, we gain important insights into the interactions between different brain regions during cognitive tasks and the deviations observed in neurological disorders. [‎15]



***Fig. 4: Brain Recording to Graph Theory Process.*** *[15]*

Related work on methods of creating graphs from brain recordings

One of the articles we studied explores changes in brain networks using EEG recordings and graph theory analysis. The study compared three groups: dementia patients, Alzheimer's patients, and a control group. During the study, the patients rested with their eyes closed while their brain activity was recorded. The researchers calculated a measure called functional connectivity, which shows how synchronized brain activity is between different electrode pairs. They used this measure to create a connectivity matrix for each participant. Next, they turned these matrices into graphs and analyzed them using various measures like typical path length, average clustering coefficient, and degree correlation. Although the article didn't specify how they chose a threshold, they mentioned they used a method called the K degree method. When they chose k = 5, this method ensured the resulting graphs were of similar sizes. The main aim of the study was to understand the structure of these brain networks and whether they differed from what's known as a "small-world" network. The findings showed that in Alzheimer's disease (AD), there was a shift away from the small-world network structure towards a more random type. These findings indicate less efficient communication between brain regions, supporting the idea of brain disconnection in AD. Contrary to expectations, patients with another condition, frontotemporal lobar degeneration (FTLD), showed changes in the opposite direction, towards a more ordered network structure. This could indicate a different underlying cause for their brain changes. This article didn't mention community detection. ‎[6]

Another article we studied on the subject explores the analysis of changes in brain networks using EEG and community detection. The article compared people with ADHD and a control group. During the test, EEG recordings were taken with eyes closed in a resting state. For each pair of electrodes, the synchronization probability measure was calculated using FSL (a new measure of general synchronization, fuzzy SL). Afterwards, the connectivity matrices were converted into weighted graphs, and the graphs between the different groups were characterized by dividing them into community structures. In the next step, the graphs were improved using a modularity maximization function. The results of the study revealed differences in the community structure of participants with ADHD and non-ADHD. ADHD patients showed more focused frontal lobe communication compared to non-ADHD individuals, indicating a lack of transfer of data from pre-processing sensory centers to high-level processing centers.[‎1]

Analysis of data set comparing information pathways in ADHD patients versus control:

This research aims to investigate differences in information pathways within brain networks between children with attention deficit hyperactivity disorder (ADHD) and healthy subjects. EEG recordings from both groups during an attentional visual task were analyzed to calculate effective connectivity using directed phase transfer entropy across different frequency bands. Group differences in connectivity between brain regions were evaluated, revealing disrupted patterns of information flow in ADHD children compared to healthy subjects. Specifically, differences were observed in posterior to anterior information flow in the theta band, and in pathways between anterior regions in the beta band. The study sheds light on the altered information flow patterns in ADHD children compared to their healthy counterparts, providing potential insights into the neural mechanisms underlying ADHD.

## **3.4** **Community Detection**

Community detection refers to identifying groups or sub-networks within a larger graph. This concept is relevant across various systems, including social, biological, and political contexts. Through community detection methods, we can algorithmically reveal the structure of networks. Understanding the structure of communities within brain networks can offer valuable insights into brain function as the brain functions in networks and not is isolated regions. A common algorithm in community detection is a modularity maximization algorithm. The algorithm aims to maximize the modularity of the network, capturing the extent to which it can be divided into distinct communities.[‎8]

Common method of Community Detection

There are a large number of methods for community detection. For each study, we need to choose the most suitable algorithm for the study that will provide us with the insights regarding the structure of the community in the networks.  
We will present several common algorithms for community detection in networks:

**Fast greedy –** The Fast-greedy algorithm, introduced by Clauset et al. in 2004, employs modularity to determine the best partitions within a network. Initially, each node is treated as its own community, and the algorithm utilizes a hierarchical clustering approach. It iteratively merges neighboring communities in a greedy manner, aiming to maximize the modularity function. This process continues until no further increase in modularity is achievable. The algorithm is based on modularity optimization principles. ‎[4]

**Louvain –** The Louvain algorithm, also known as Multi-level, introduced by Blondel et al. in 2008, shares similarities with the Fast-greedy algorithm as it employs modularity for partition optimization. However, it focuses on identifying hierarchical structures within large networks. The algorithm iteratively swaps nodes between communities, evaluating modularity changes until no further improvement is possible. Subsequently, it collapses communities into latent nodes and establishes edge weights between observed and latent nodes, creating a multi-level structure. The final outcome of the Louvain algorithm is influenced by the node order, rendering it non-deterministic. ‎[4]

**Edge betweenness –** The Edge Betweenness algorithm, introduced by Girvan & Newman in 2002, was among the pioneering methods for community detection in networks. It identifies edges that frequently connect different nodes, known as edge betweenness, based on betweenness centrality. Initially, edge betweenness is computed for the entire network, and the edge with the highest value is removed. Subsequently, affected edges have their betweenness recalculated, and this process repeats iteratively until no edges remain. Although slower compared to other algorithms, modularity is employed to determine the optimal cut-off point. This algorithm utilizes an approach based on edge centrality detection. ‎[4]

**Label propagation –** The Label Propagation algorithm, introduced by Raghavan et al. in 2007, starts by giving each node a distinct label. Nodes then adopt the label shared by the majority of their neighbors, with random resolution of ties. This iterative process continues until all nodes share the label of the majority of their neighbors. The underlying idea is that nodes will reach a consensus over time. Similar to the Louvain algorithm, this method is non-deterministic and yields varying results with each execution. This algorithm employs an approach based on topological closeness. ‎[4]

# **Expected Achievements**

In our project, we will concentrate on examining brain recordings from two distinct groups: individuals diagnosed with ADHD and a control group. Our objectives are outlined as follows:

1. Our system will facilitate users in visualizing and exporting graphs generated from EEG recordings.
2. We intend to showcase the local networks delineated within the graph using an appropriate community detection algorithm and highlight the variances between the groups within these networks.

# **Engineering Process**

## **5.1 Process**

In order to proceed with our project, our initial step involved acquiring an understanding of attention deficit hyperactivity disorder (ADHD), including its defining characteristics and impact on individuals. Subsequently, we delved into Electroencephalography (EEG), a diagnostic method utilized for assessing brain activity through the recording of electrical signals. Our exploration encompassed the distinct frequency waves associated with EEG tests, elucidating their occurrence and significance.

Advancing further, our focus shifted towards the exploration of graph theory as a means of network representation. We discerned that graph theory furnishes a robust mathematical framework for the analysis of networks and their information dynamics. Our inquiry extended to the calculation of arc weights, as well as the identification and comprehension of both global and local indices pertinent to graph analysis. Within the context of our project, we familiarized ourselves with the concept of community detection within networks, along with the diverse methodologies employed for community detection within graph structures.

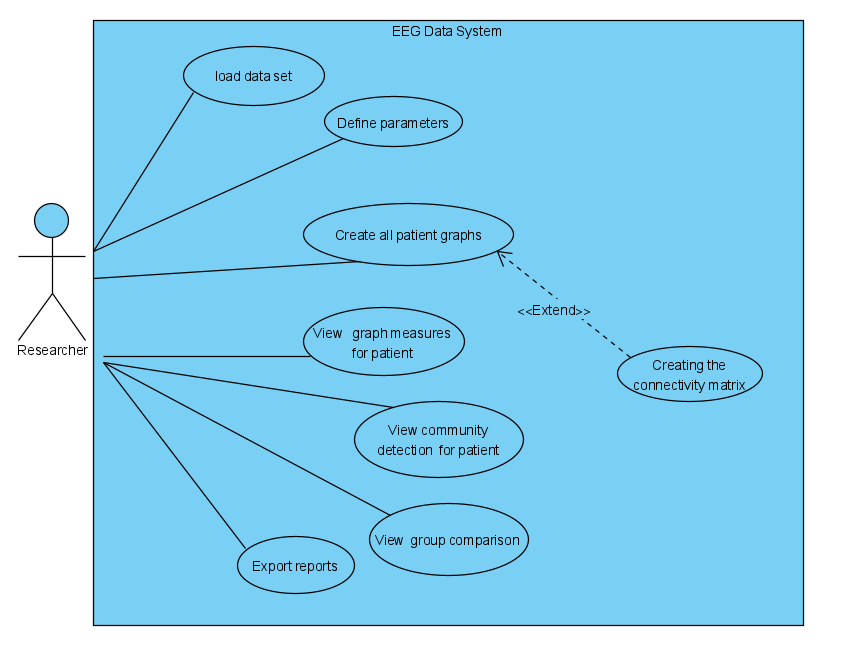
We intend to analyze a dataset comprising brain recordings from two distinct groups: children diagnosed with attention deficit hyperactivity disorder (ADHD) and a control group of healthy children. Our dataset consists of 61 children with ADHD and 60 typically developing children aged between 7 and 12, encompassing both genders. The EEG recordings were conducted utilizing 19 channels at a frequency of 128 Hz.

During the EEG recordings, participants were engaged in a visual attention task where they were instructed to count cartoon characters presented in images. The number of characters varied randomly between 5 and 16, with the images designed to facilitate clear viewing and accurate counting.

Subsequently, we delineate the recording range for analysis, transforming the recordings into graph representations. This process entails selecting an appropriate methodology for determining arc weights, establishing suitable thresholds, and identifying global and local graph characteristics for exploration. Additionally, we opt for a specific community detection approach within the graph.

## **Product**

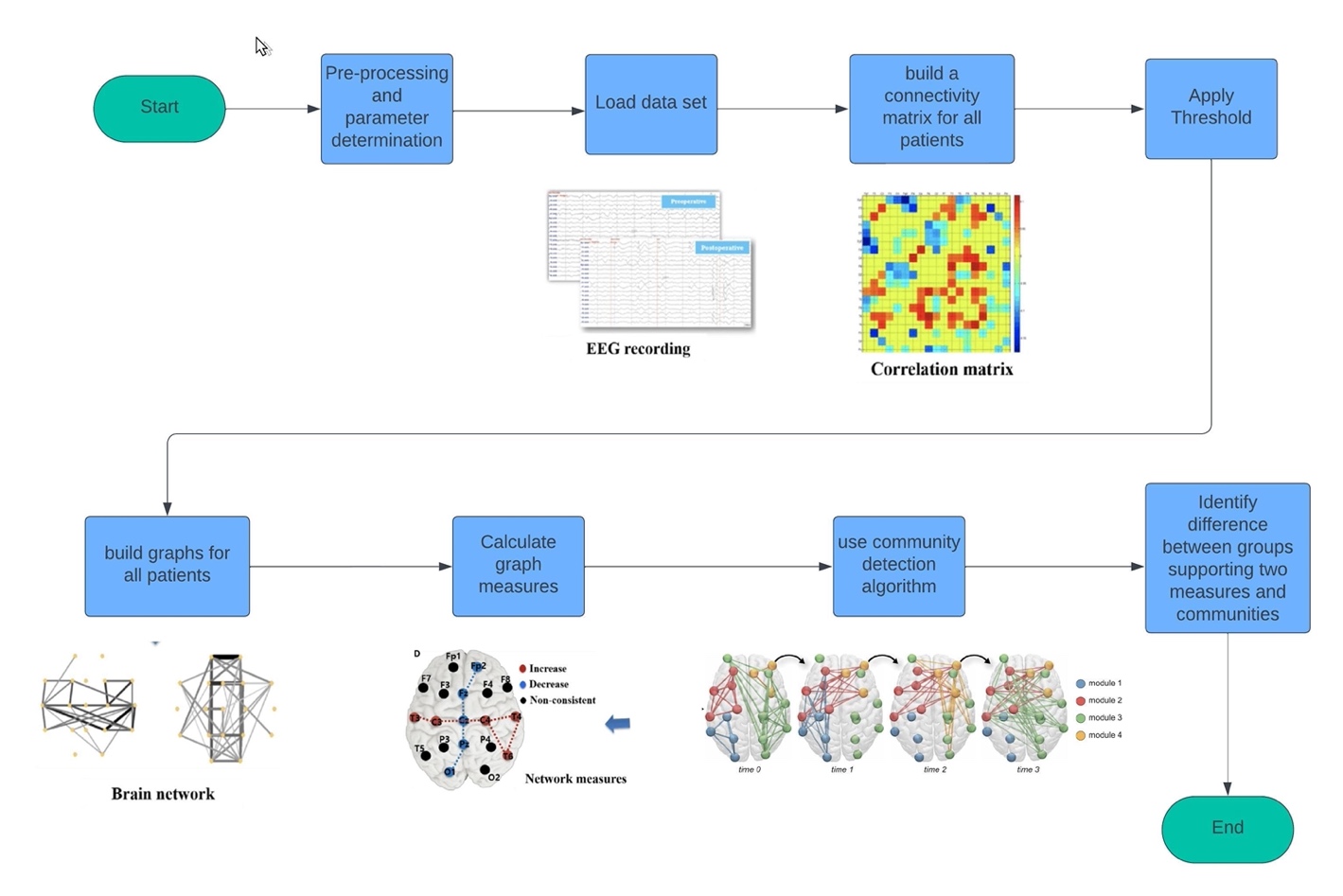
### **5.2.1 use case diagram**



***Figure 5. Use case diagram.***

Our system facilitates seamless data loading and visualization of brain recordings in graphical format. In addition, the system allows setting different parameters, such as, time range of the recording. The system can generate comprehensive graphs for each individual, which encompasses the ability to view the different measures, community detection, and conduct a thorough comparison between the two graphs.

### **5.2.2 Data Process**



***A***

***B***

***C***

***D***

***E***

***F***

***G***

***H***

***Figure 6. Flow chart diagram.***

**A** - **Pre-processing and parameter determination –**

There are 4 methods of determining weight of wave signals, Pearson correlation coefficient, Spearman rank order correlation, Kendall rank order correlation, and Mutual information.

We chose to utilize the Pearson’s correlation coefficient, which is the most widely used technique to determine the correlation of signals between a pair of nodes. This method is widely used for several reasons, due to its simplicity, robustness, and interpretability. It offers a straightforward measure of the strength and direction of linear relationships between two variables, ranging from -1 to 1, making it easily understandable and communicable. Its resistance to outliers ensures its applicability across diverse datasets, while its standardized nature allows for comparisons between variables with different units. Moreover, its extensive use in various fields, along with its availability in statistical software, enhances its acceptance and accessibility. Despite its assumption of linearity, which may not always hold, the Pearson correlation coefficient remains a widely used and reliable tool for assessing relationships between variables in research and analysis. ‎[10]

The Pearson correlation coefficient (r) quantifies the strength and direction of the linear relationship between two variables. It always falls within the range of -1 to +1. A value closer to 1 or -1 indicates a stronger linear correlation, while a value closer to 0 suggests a weaker correlation. Variance measures the extent to which individual values in a dataset deviate from the mean value. A positive r value signifies a positive correlation, meaning that as one variable increases, the other variable tends to increase as well. Conversely, a negative r value indicates a negative correlation, where an increase in one variable corresponds to a decrease in the other. The coefficient of determination (r^2) represents the proportion of variance in one variable that can be explained by variance in the other variable. However, correlation does not imply causation, meaning that even if two variables are strongly correlated, it does not necessarily mean that changes in one variable cause changes in the other. ‎[11]

***Formula 1. Pearson’s formula ‎ [11]***

The Pearson correlation coefficient (PCC) based channel selection method aims to identify and eliminate redundant EEG channels in a dataset by evaluating the correlation between channel pairs during motor imagery (MI) tasks. This method assumes that channels related to motor imagery exhibit consistent correlations across multiple trials. The PCC is calculated for each channel pair, and channels with high correlations are grouped together. By selecting the most frequently correlated channels, redundant channels are removed, resulting in a reduced dataset. This reduction in dataset size improves computational efficiency for feature extraction and enhances classification accuracy for MI Brain-Computer Interface (BCI) systems. ‎[9]

Method of threshold determination:

In our meta-analysis of several papers, we have identified two primary approaches used to determine thresholds. One study technique employs a method to determine the threshold for forming graphs based on synchronization likelihood (SL) values, aiming to ensure consistency in the resulting network graphs across different groups. By selecting thresholds that yield equal average degrees for both groups, any remaining differences between the networks are more likely to reflect genuine disparities in network organization. The method involves examining various thresholds to ensure robustness. [‎6]

Another statistical based method focuses on synchronization of matrices. The initial step in applying graph theoretical analysis to synchronization matrices involves converting the N × N synchronization matrix into a binary graph. This binary graph represents connections between elements, with edges either existing or not existing based on a chosen threshold (T). As there's no single optimal threshold, a range of values (0.01 < T < 0.05, with increments of 0.001) was explored. For each threshold value, if the synchronization likelihood (SL) between a pair of channels exceeds T, an edge is established between them; otherwise, no edge is formed. This process was repeated for the entire range of threshold values to ensure comprehensive analysis. ‎[17]

There exist several thresholds and methods for determination. Our goal is to assess and contrast various methods to identify the most suitable threshold method for analysis in this project. This aspect will be explored and finalized in part B of our project, currently designated as "To Be Determined".

**B - loading data set –**

A study that utilized the investigated dataset categorized the time series as follows: Each subject's time series were partitioned into segments lasting 8 seconds each, resulting in a total of 1024 samples. The quantity of segments varied among subjects due to variations in the timing task, ranging from 50 seconds as the minimum for a subject in the control group to 285 seconds as the maximum for a subject with ADHD. On average, the control group had 13.18 segments (with a standard deviation of 3.15), while the ADHD group had 16.14 segments (with a standard deviation of 6.42). [‎7]

**C - build connectivity matrix for all patients –**

The connectivity matrix (CM) serves as a fundamental tool in understanding the intricate networks of the brain, particularly in relation to structural, functional, and effective connectivity. It encompasses a full square matrix with N^2 elements, where N represents the number of neurons, brain regions, or assemblies under consideration.

Three main types of connectivity matrices are discussed: functional, effective, and structural. Functional connectivity pertains to correlations between time series from different sources, without any underlying causal model. Effective connectivity identifies direct influences exerted by one neuronal system on another, relying on structural connections within a network model. Structural connectivity considers the physical, synaptically mediated connections among neurons or assemblies. ‎[3] Using the Pearson’s coefficient we can calculate the connectivity matrix for each patient.

**D- Apply Threshold**

As we mentioned above, the selection of the threshold will be determined in part B of the project when we examine possible thresholds according to the results from the connectivity matrix.

**E- Build graphs for all patients**

After selecting the threshold and calculating it, we will accordingly build a graph for each patient whose nodes in the graph represent the locations of the electrodes, while the edges indicate functional connectivity between these locations. After constructing the graph, we can gain important insights into the interactions between different brain regions during cognitive tasks for each patient.

**F- Calculate graph measures-**

The calculated measures are:

Local Clustering Coefficient -

The Local Clustering Coefficient (LCC) serves as a metric for assessing network segregation by gauging the extent to which neighboring nodes form cohesive groups or cliques. It is computed for each node (i) by comparing the sum of geometric means of all existing weighted triangles to the total number of potential triangles. Essentially, LCC reflects the effectiveness of communication within the node's cluster, with higher values indicating greater local efficiency in information transfer. ‎[2]

***Formula 2. Local Clustering Coefficient. (19)***

Average Degree -

Average degree is a fundamental metric in network analysis that calculates the average number of edges per node within a graph. This value provides insight into the level of connectivity within the network. To compute the average degree, the total number of edges in the network is divided by the total number of nodes. [‎19]

***Formula 3. Average degree formula. ‎[19]***

Global efficiency -

Global Efficiency (GE) serves as a comprehensive metric in network analysis, assessing the efficiency of information exchange across the entire network. It is calculated as the average of the inverses of the shortest paths between all pairs of nodes within the network.

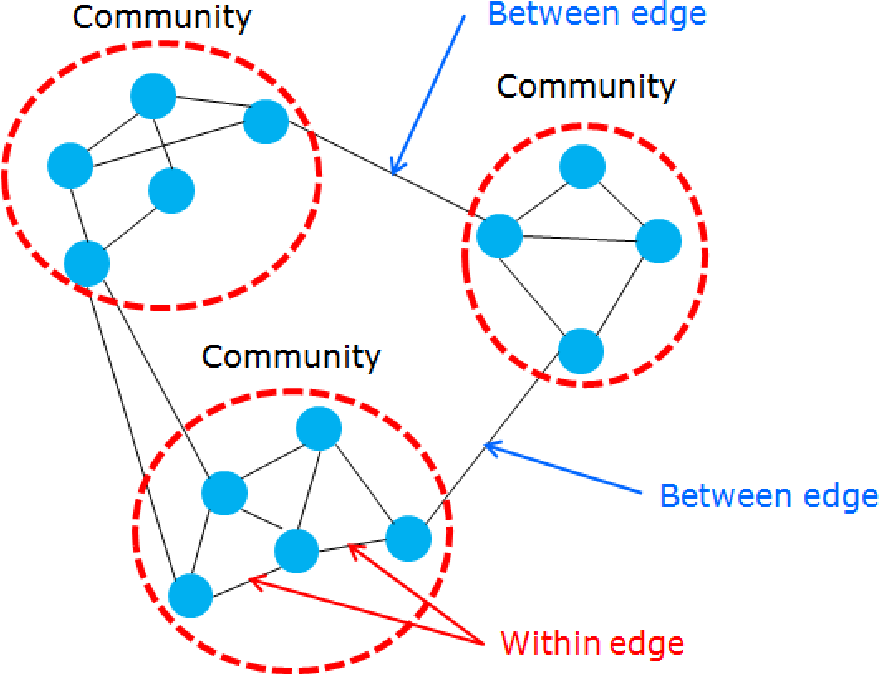
GE offers insights into the effectiveness of communication among nodes, with higher values indicating more efficient information exchange. Networks with high global efficiency are characterized by short communication paths between nodes, reflecting strong global integration.

Overall, GE provides a quantitative measure of the network's ability to facilitate communication among its nodes, highlighting its capacity for global integration. [‎16]

***Formula 4. Global efficiency formula ‎[12]***

**G – use community detection algorithm**

The fast greedy algorithm presents an efficient method for identifying communities within networks by maximizing modularity. Initially, it constructs a subnetwork consisting solely of connections among densely linked nodes. Subsequently, the algorithm iteratively selects random links that enhance the subnetwork's modularity and incorporates them. This iterative procedure continues as long as modularity enhancements are observed. Ultimately, communities are delineated based on the connected components present within the subnetwork. ‎[20]



***Figure 7: Community detection. [‎21]***

Following the complete the fast greedy algorithm, we measure modularity of the detected communities.

The modularity of a graph with respect to some division (or vertex types) measures how good the division is, or how separated are the different vertex types from each other. It defined as:

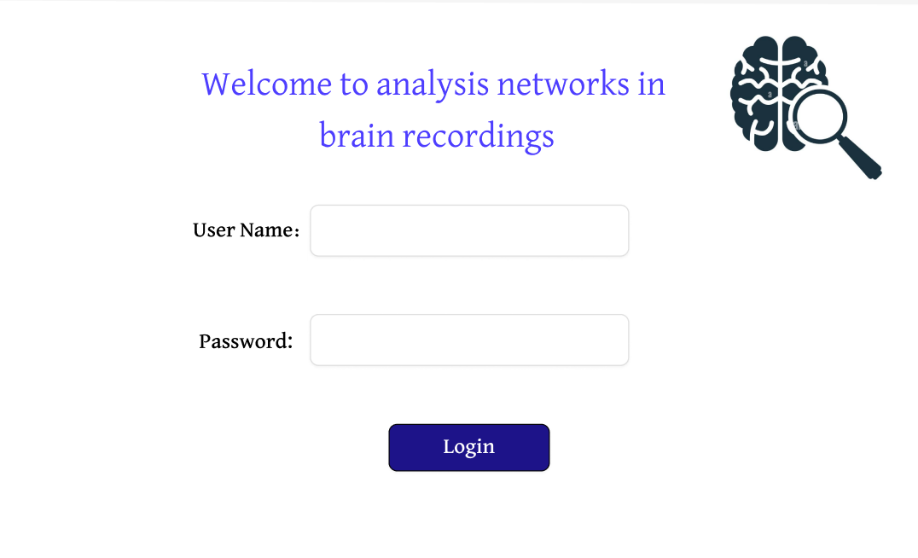
***Formula 5. Modularity formula***

Here m is the number of edges, is the element of adjacency matrix in row *i* and column *j*, is the degree of *i*, is the degree of *j*, is the type (or component) of *i*, that of *j*, the sum goes over all *i* and *j* pairs of vertices, and is 1 if and 0 otherwise. [‎22]

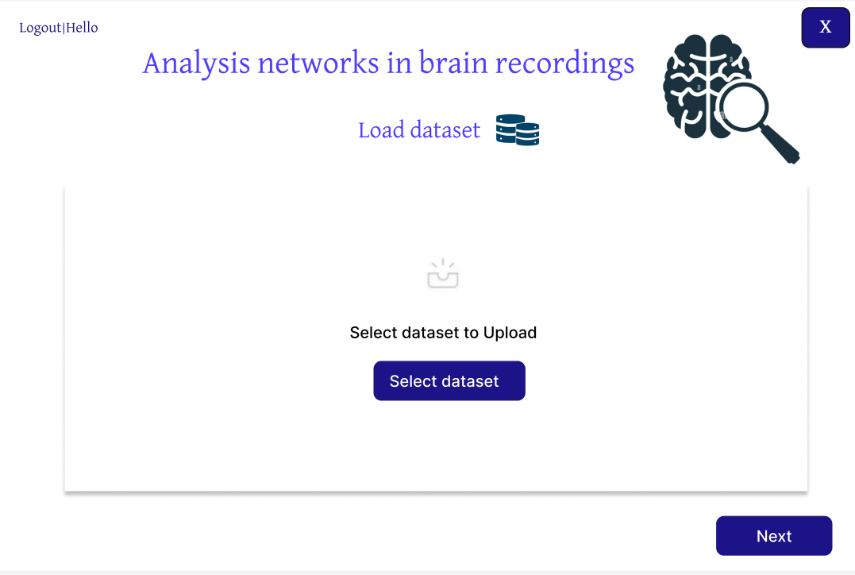
**H – Identifying differences between groups supporting two measures and communities –**

In this stage, comparison between graphs of ADHD to control cases is performed. Analysis of community detection and the graph measures is done to compare differences between two groups or between individuals.

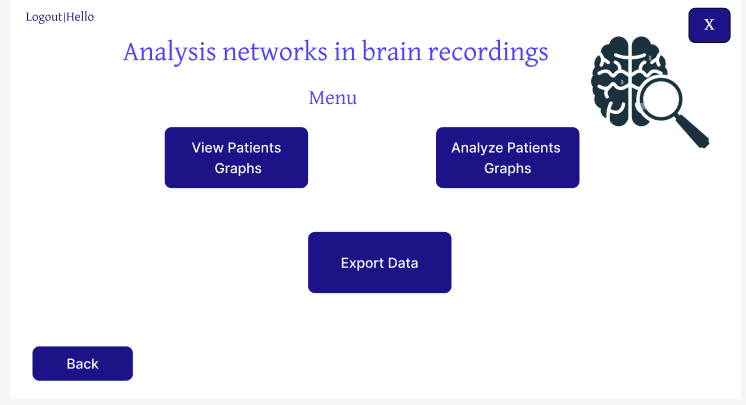
### **5.2.3 Prototype**

****

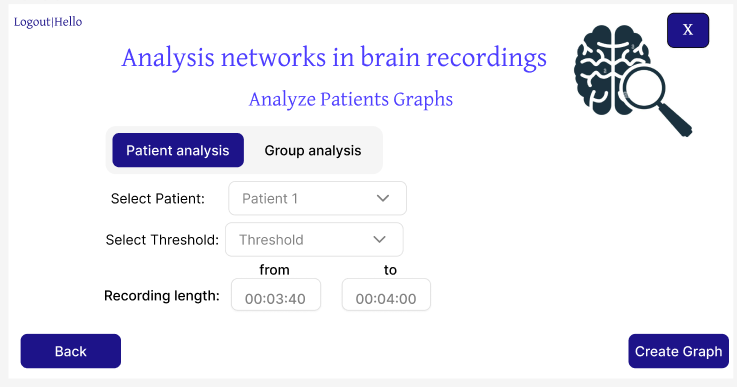
login screen page. The exporter can login to the system.



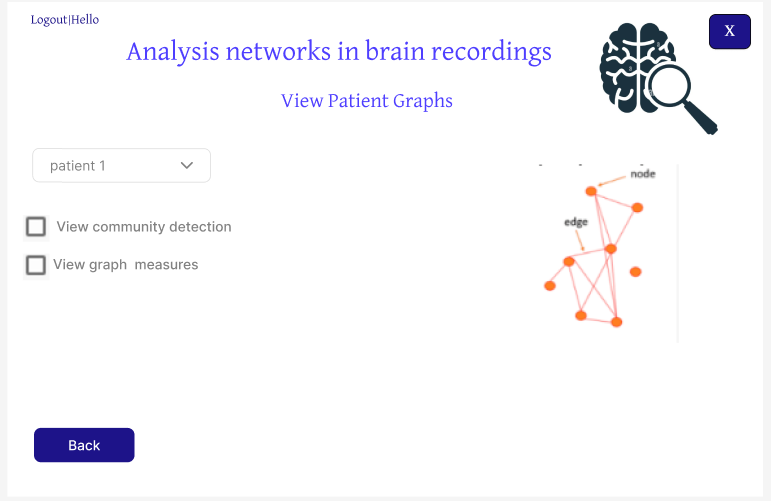
Dataset screen page. In this page the exporter can upload the brain recordings dataset.



Menu screen page. In this page the exporter can choose to view patient's graphs, analyze patients graphs and export the data.



Analyze patient's screen. In this screen, the exporter selects if he wants to see patient analysis or group analysis (ADHD and healthy graphs) set graph thresholds and choose the recording duration they want to focus on.



View patient graphs screen. In this screen the exporter selects the patient and views their graph. Additionally, the exporter has the option to view community detection and graph measures within the graph.

# **Verification plan**

|  |  |  |
| --- | --- | --- |
| Case | Test Case | Expected Result |
| 1 | Insert same signal to connectivity matrix | Because we use Pearson correlation calculation the expected result of the same signal is 1. |
| 2 | Insert different signals to connectivity matrix | Because we use Pearson correlation calculation the expected result of different signals is between 0-1. |
| 3 | Apply threshold | It is observed that the results of the chosen threshold will be the same for articles with the same data set and the same threshold. |
| 4 | Build graph for all patients. | It was observed that the graphs of the patients would be similar to the graphs in the articles with the same data set. |
| 5 | Calculate graph measures | It will be observed that if we take a simple graph and calculate all the measures on it or for a graph whose results we know we will get a correct calculation of the measures. |
| 6 | Using community detection algorithm | We will observe that if we take a graph whose community modules are known and put it into our community detection algorithm we will expect to get the same modules in the graph. |

As there is an existing paper that utilizes the same dataset as our project, we can compare our findings with those reported in the original research paper to identify any similarities.

Analysis of Graph Measures Between Groups

We are interested to understand the differences between the ADHD and control groups. We'll do this by calculating the average graph measures for each group and then comparing them.

# **Reference:**

1. Ahmadlou M, Adeli H. Functional community analysis of brain: A new approach for EEG-based investigation of the brain pathology. NeuroImage. 2011 Sep 15;58(2):401–8.
2. Al-Shargie F, Hassanin O, Tariq U, Al-Nashash H. EEG-Based Semantic Vigilance Level Classification Using Directed Connectivity Patterns and Graph Theory Analysis. IEEE Access. 2020 Jun 23;PP:115941–56.
3. Boschi A, Brofiga M, Massobrio P. Thresholding Functional Connectivity Matrices to Recover the Topological Properties of Large-Scale Neuronal Networks. Front Neurosci. 2021 Aug 16;15:705103.
4. Christensen AP, Garrido LE, Guerra-Peña K, Golino H. Comparing community detection algorithms in psychological data: A Monte Carlo simulation [Internet]. OSF; 2020 [cited 2024 Apr 6]. Available from: <https://osf.io/hz89e>
5. Contributors WE. WebMD. [cited 2024 Apr 6]. What to Know About Gamma Brain Waves. Available from: https://www.webmd.com/brain/what-to-know-about-gamma-brain-wave
6. de Haan W, Pijnenburg YA, Strijers RL, van der Made Y, van der Flier WM, Scheltens P, et al. Functional neural network analysis in frontotemporal dementia and Alzheimer’s disease using EEG and graph theory. BMC Neurosci. 2009 Aug 21;10:101.
7. Ekhlasi A, Nasrabadi AM, Mohammadi MR. Direction of information flow between brain regions in ADHD and healthy children based on EEG by using directed phase transfer entropy. Cogn Neurodyn. 2021 Dec;15(6):975–86
8. Garcia JO, Ashourvan A, Muldoon SF, Vettel JM, Bassett DS. Applications of community detection techniques to brain graphs: Algorithmic considerations and implications for neural function. Proc IEEE Inst Electr Electron Eng. 2018 May;106(5):846–67.
9. Pawan, Dhiman R. Electroencephalogram channel selection based on pearson correlation coefficient for motor imagery-brain-computer interface. Meas Sens. 2023 Feb 1;25:100616
10. Pernet CR, Wilcox R, Rousselet GA. Robust Correlation Analyses: False Positive and Power Validation Using a New Open Source Matlab Toolbox. Front Psychol. 2013 Jan 10;3:606.
11. Turney S. Scribbr. 2022 [cited 2024 Apr 6]. Pearson Correlation Coefficient (r) | Guide & Examples. Available from: https://www.scribbr.com/statistics/pearson-correlation-coefficient/
12. Wang J, Zuo X, He Y. Graph-Based Network Analysis of Resting-State Functional MRI. Front Syst Neurosci. 2010 Jun 7;4:16.
13. What is ADHD? [Internet]. [cited 2024 Apr 6]. Available from: https://www.psychiatry.org:443/patients-families/adhd/what-is-adhd
14. Doing an EEG – The Patrick Wild Centre [Internet]. [cited 2024 Apr 6]. Available from: https://patrickwildcentre.com/doing-an-eeg/
15. Labs S. A Primer on the Application of Graph Theory to EEG [Internet]. Sapien Labs | Neuroscience | Human Brain Diversity Project. 2019 [cited 2024 Apr 6]. Available from: https://sapienlabs.org/graph-theory-in-eeg/
16. Graph Measures [Internet]. BRAPH. [cited 2024 Apr 6]. Available from: http://braph.org/braph-1-0/manual/graph-measures/
17. Small-World Networks and Functional Connectivity in Alzheimer’s Disease | Cerebral Cortex | Oxford Academic [Internet]. [cited 2024 Apr 6]. Available from: https://academic.oup.com/cercor/article/17/1/92/283419?login=false
18. Clustering Coefficient in Graph Theory - GeeksforGeeks [Internet]. [cited 2024 Apr 6]. Available from: https://www.geeksforgeeks.org/clustering-coefficient-graph-theory/
19. Jilbert OL and I. 2.7 Average Degree | Social Networks: An Introduction [Internet]. [cited 2024 Apr 6]. Available from: https://bookdown.org/omarlizardo/\_main/2-7-average-degree.html
20. Community Detection [Internet]. [cited 2024 Apr 6]. Available from: https://noesis.ikor.org/wiki/algorithms/community-detection
21. Community detection using efficient modularity optimization method: labelmod with single and multi-layer graphs | Semantic Scholar [Internet]. [cited 2024 Apr 6]. Available from: https://www.semanticscholar.org/paper/Community-detection-using-efficient-modularity-with-Bang/813dc781d5d44f56ea1fc1454d8b441f2f6069a5
22. igraph R manual pages [Internet]. [cited 2024 Apr 6]. Available from: https://igraph.org/r/doc/modularity.igraph.html

**Git Link:**

<https://github.com/KoralBiton18/Capstone_Project_Inception.git>