Software Engineering Department
Braude College

Capstone Project Phase A

EEG Classification Using Text Compression

25-1-R-2

By:- Mohammad khateeb
Jad taha

Advisors:- Dr. Samah Idrees Ghazawi
Dr. Anat Dahan

Link to github:-
https://github.com/mhmdkh1905/EEG-recordings.git

Contents

Abstract

Attention-Deficit/Hyperactivity Disorder (ADHD) is a condition that affects attention and behavior, making accurate diagnosis essential. This study uses electroencephalography (EEG), a non-invasive measure of brain activity, to explore differences between children with ADHD and typically developing controls. We will apply a text compression technique called Normalized Compression Distance (NCD) or similar techniques in using text compression algorithms to classify EEG recordings. After cleaning and normalizing the data, we will calculate similarity scores between EEG signals to identify patterns unique to each group. By analyzing these patterns across different brain regions, this method aims to improve the accuracy of ADHD classification and contribute to better diagnostic tools.

1. Introduction

Our project investigates the application of text compression techniques to EEG classification by comprehensively reviewing existing methods and selecting the most relevant and accurate approaches.

Text compression algorithms, designed to identify and exploit duplicates within data sequences, have shown promise in various classification tasks. Rather than developing a new method, this research explores existing studies, analyzes the methodologies applied to classification problems, and identifies those most suitable for EEG signal classification. By leveraging proven techniques, we aim to assess their performance and applicability in this domain.

In the next section, we provide background information, including an introduction to EEG signals, ADHD, text compression techniques, and related work, as well as details about the dataset we intend to use and relevant studies. Finally, we will present our approach for integrating and evaluating these existing methods within the context of EEG classification.

2. Background and Related Work

2.1. Introduction to EEG

We introduce the problem of understanding and analyzing brain activity through non-invasive methods and present the foundational principles of electroencephalography (EEG). EEG is a widely used diagnostic tool that records the brain's electrical activity by placing electrodes on the scalp. Its significance lies in its ability to detect and monitor neurological disorders, offering insights into both normal and pathological brain functions [3].

Today, with the increasing prevalence of neurological conditions and the growing demand for advanced diagnostic techniques, EEG plays a crucial role in clinical and research settings. This technology utilizes the principles of differential amplification to record electrical signals generated by neuronal activity, providing a non-invasive and real-time view of brain function that can be seen in Fig. 1. EEG recordings are often used to identify patterns such as normal rhythms (e.g., alpha waves) and pathological discharges (e.g., epileptiform activity), enabling clinicians to make accurate diagnoses [17].

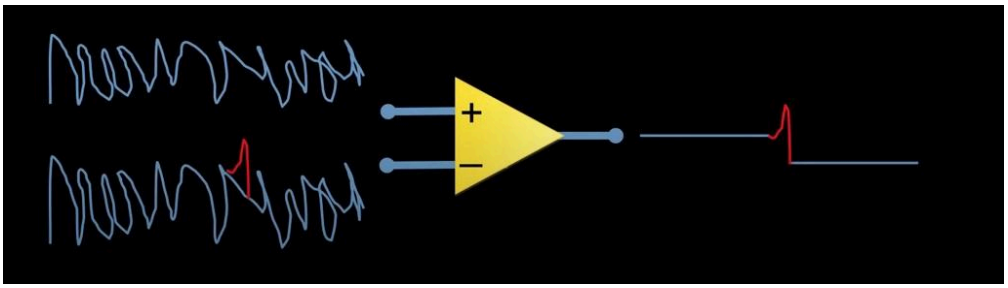

Fig. 1. Amplifiers that have two inputs of EEG signals, and the output is the difference between them.

EEG systems rely on standardized electrode placement protocols, such as the international 10–20 system, to ensure consistency and reliability in recordings. This system uses key anatomical landmarks such as the nasion (bridge of the nose), inion (bony prominence at the back of the head), and preauricular points (in front of the ears) to guide electrode positioning. The electrodes are labeled according to their cortical region: F for frontal, T for temporal, C for central, P for parietal, and O for occipital. Odd numbers indicate the left hemisphere, even numbers indicate the right hemisphere, and Z is used for electrodes along the midline, which is shown in Fig. 2. For example, F3 and F4 correspond to the left and right frontal regions, respectively, while Cz represents the central midline position [12].

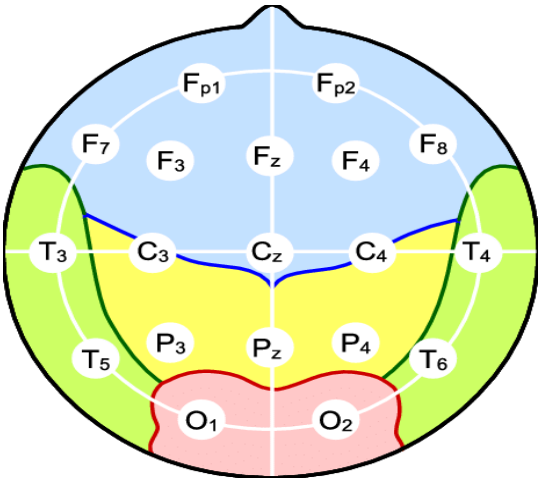

Fig. 2. Visual representation of the 10-20 system, showing electrode positions on a human head.

In conclusion, EEG remains an essential tool in neuroscience and clinical medicine, providing invaluable insights into brain function and aiding in the diagnosis of neurological disorders. Its non-invasive nature, coupled with its ability to capture real-time data, makes it a cornerstone of modern diagnostics. By understanding the principles and applications of EEG, researchers, and clinicians can continue to advance the field, improving outcomes for patients with neurological conditions.

2.2. Introduction to ADHD

Attention-Deficit/Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder characterized by persistent patterns of inattention, hyperactivity, and impulsivity that interfere with functioning or development. It affects approximately 5% of children and adolescents worldwide and often persists into adulthood. The etiology of ADHD is multifactorial, involving genetic, environmental, and neurobiological factors [7].

Diagnosing ADHD involves a comprehensive evaluation, as there is no single test to confirm the condition. Clinicians typically assess behavioral and mental development, gather feedback from parents and teachers, and adopt structured assessment measures. This process helps differentiate ADHD from other conditions with similar symptoms, such as sleep disorders, anxiety, or depression [18].

Electroencephalography (EEG) has been explored as a tool to identify biomarkers associated with ADHD. Research indicates that individuals with ADHD often exhibit distinct EEG patterns, such as increased theta wave activity and altered connectivity. These findings highlight EEG's potential to provide valuable insights into the neurophysiological underpinnings of ADHD, aiding in its diagnosis and understanding [1].

2.3. Dataset Overview

Understanding the characteristics of the dataset is essential for contextualizing the scope and methodology of our research. In this section, we provide detailed information about the EEG dataset in our study, including its structure, participant demographics.

2.3.1. Description of the Dataset

In our work, we leveraged a dataset consisting of EEG recordings from 60 children diagnosed with ADHD and 60 healthy controls, all aged between 7 and 12 years. The ADHD children were diagnosed based on DSM-IV criteria by experienced psychiatrists, ensuring clinical accuracy, and had been treated with Ritalin for up to six months before the recordings, allowing us to account for medication effects. The control group was screened for psychiatric, neurological, or behavioral disorders to ensure a reliable comparison. EEG data were collected using the standardized 10–20 system across 19 channels (e.g., Fz, Cz, Pz) at a sampling frequency of 128 Hz. Participants engaged in a visual attention task, designed to stimulate cognitive processing, by counting cartoon characters in randomly displayed images, an example of these images is shown in Fig. 3. This structured protocol provided a robust dataset for comparing neural dynamics and connectivity patterns between ADHD and typically developing (TD) children, forming the basis for our investigations [19].

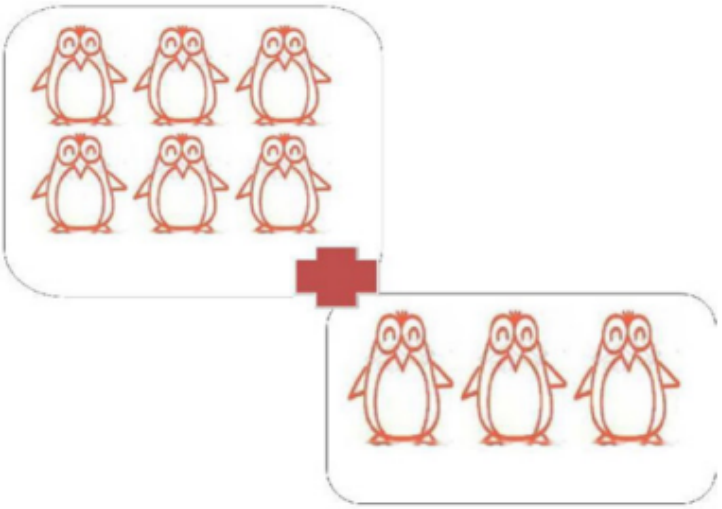

Fig. 3. An example of the images shown to the children during the EEG signals recording.

This EEG dataset is an important tool for studying the brain activity of children with ADHD compared to healthy children, especially in tasks that test visual attention, which is often a challenge for kids with ADHD. The dataset is well-designed, with accurate diagnoses, carefully matched participants, and standard EEG recording methods. The attention task is engaging and adjusts to each child's response speed, providing a clear picture of how their brains work during these tasks. This makes the dataset useful for finding patterns in brain activity that could help diagnose ADHD, understand how treatments like Ritalin affect the brain, and create better tools for identifying and managing ADHD. Its ability to show clear differences in brain activity during specific tasks makes it a powerful resource for research and practical applications.

We will use the raw EEG data from this dataset to preserve all details and patterns in brain activity. This allows us to analyze the signals directly and apply advanced methods to better understand the differences in ADHD brain activity, ensuring more accurate and reliable results. Fig. 4 below presents an example of the raw EEG data, illustrating the recorded brain signals before preprocessing.

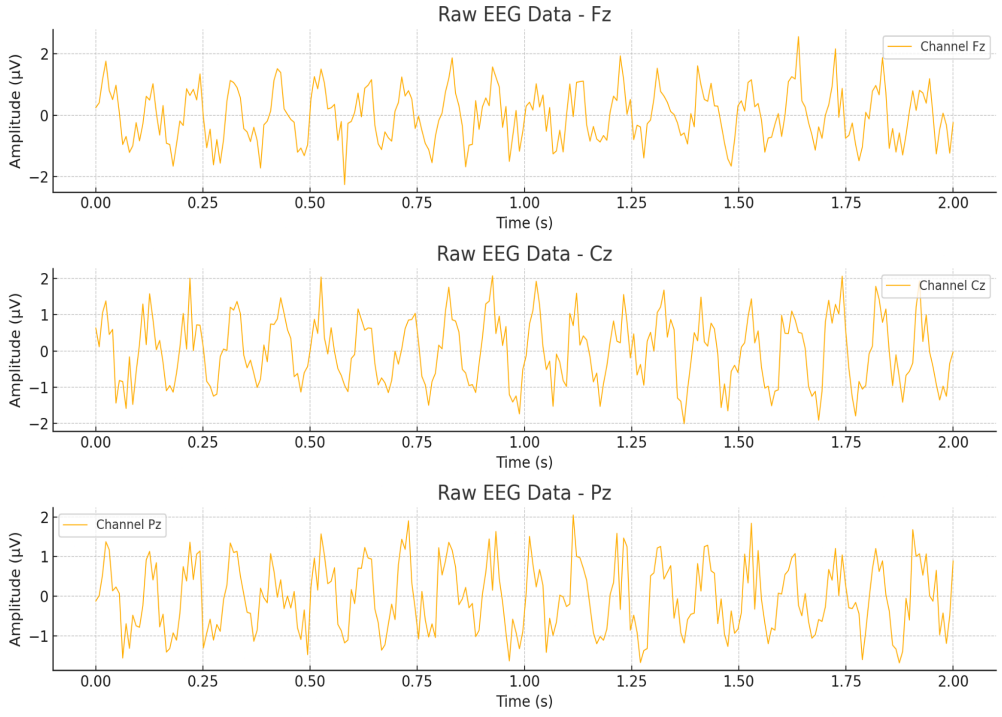

Fig. 4. Example of raw EEG data recorded from three channels (Fz, Cz, and Pz) using the 10–20 system. The x-axis represents time (seconds), and the y-axis represents amplitude (microvolts). These signals reflect brain activity before any preprocessing

or filtering, capturing the natural variability in neural dynamics during the visual attention task.

2.3.2. Prior Work Using the Dataset

After reviewing academic literature related to the dataset we plan to use, we found several studies demonstrating the potential of EEG data in identifying brain activity patterns associated with ADHD. Collectively, these studies highlight the versatility and promise of various EEG analysis methods, from connectivity measures to non-linear feature extraction, in advancing ADHD diagnosis.

One study specifically examined differences in brain connectivity between children with ADHD and typically developing (TD) children during a visual attention task. The study used advanced methods, including the nCREANN approach to analyze both linear and nonlinear brain connectivity, and the direct Directed Transfer Function (dDTF) to measure spectral connectivity. These methods achieved an impressive classification accuracy of 99%, showing their effectiveness in distinguishing ADHD from TD groups. This research emphasizes the importance of integrating both linear and nonlinear measures to enhance the diagnostic power of EEG-based tools [16].

Building on the importance of connectivity measures, another study explores how information flows between brain regions in children with ADHD compared to TD children. By using the directed Phase Transfer Entropy (dPTE) method, this study examines the strength and direction of brain connectivity across EEG frequency bands, such as delta, theta, and beta. The findings reveal significant differences in connectivity patterns: children with ADHD exhibit disrupted communication from the back (posterior) to the front (anterior) of the brain, particularly in the theta band, and reduced connectivity in the front (anterior) in the beta band, as we can found it in Fig.5 . These results underscore the value of frequency-specific connectivity analyses in identifying critical biomarkers of ADHD [8].

Extending beyond connectivity patterns, another study delves into the classification of ADHD and healthy children by analyzing nonlinear features of EEG signals. This research explores complex patterns in brain activity using measures such as fractal dimension, approximate entropy, and the Lyapunov exponent. To identify the most relevant features, the study employs feature selection methods like Double Input Symmetrical Relevance (DISR) and Minimum Redundancy Maximum Relevance (mRMR). These features are then used as inputs to a Multi-Layer Perceptron (MLP) neural network, achieving classification accuracies of 93.65% with DISR and 92.28% with mRMR. These findings demonstrate the power of non-linear features in complementing connectivity-based approaches [13].

Further extending the scope of non-linear analyses, another study focuses on attention continuity as a distinguishing feature for classifying ADHD and healthy children. This study examines features like the Lyapunov exponent and various fractal dimensions

(Higuchi, Katz, and Sevcik) extracted from EEG data. Using these features, a Multi-Layer Perceptron (MLP) neural network achieves a remarkable accuracy of 96.7%, with the best results obtained from EEG data recorded from the frontal region of the brain. These results highlight the crucial role of the frontal lobe in attention-related processes and reinforce the importance of combining region-specific analyses with non-linear features for a comprehensive understanding of ADHD [2].

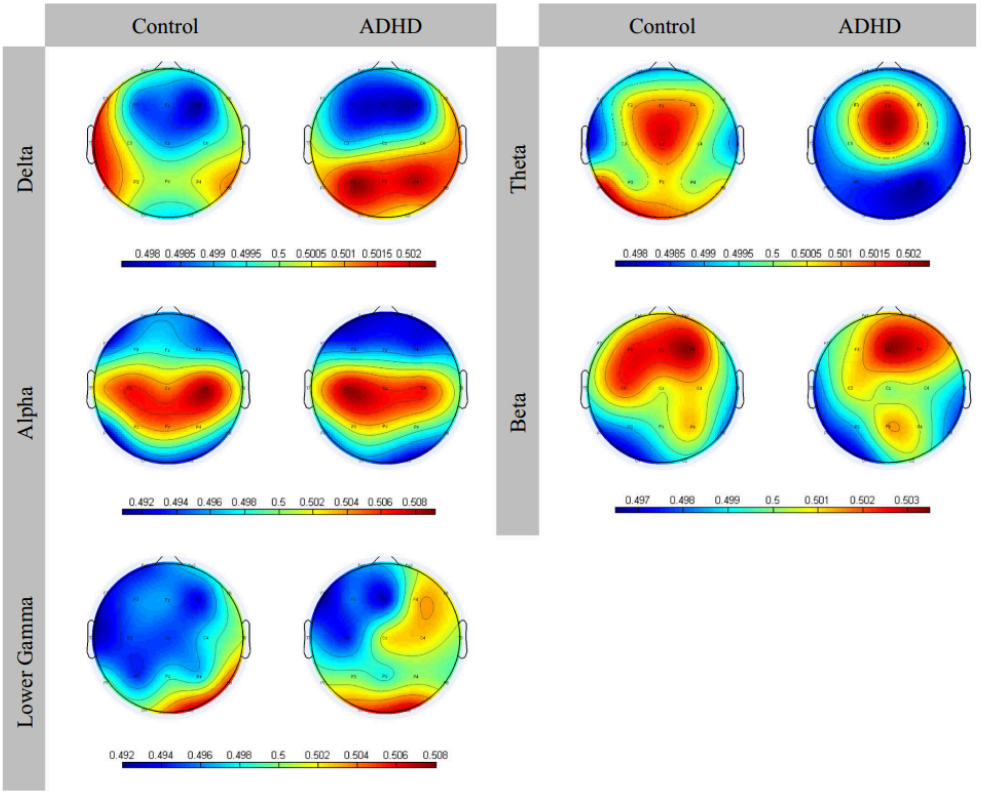

Fig. 5. The mean dPTE (Directional Phase Transfer Entropy) for each brain region is shown as a color-coded map. Red indicates areas where information is flowing out, while blue shows areas where information is flowing in. In this map, differences are noticeable between Control and ADHD subjects. Specifically, in the Delta and Theta frequency bands, there are changes in the posterior (back) region, while in the Beta frequency band, changes appear in the anterior (front) region.

2.3.3. Key Findings

The reviewed articles provide useful insights into how EEG analysis can help understand and diagnose ADHD. Key findings show that altered brain connectivity, such as disruptions in posterior-to-anterior and anterior connectivity patterns, is significant in ADHD children, especially in the theta and beta frequency bands. Nonlinear features like fractal dimensions, approximate entropy, and Lyapunov exponents are effective in capturing the

chaotic nature of EEG signals in ADHD, with classification accuracies reaching up to 96.7% using advanced neural networks like Multi-Layer Perceptron (MLP). The studies also highlight the importance of combining linear and nonlinear connectivity measures, focusing on specific brain regions like the frontal lobe, and using innovative features such as attention continuity to improve the diagnostic power of EEG systems. They further show that using feature selection methods like Double Input Symmetrical Relevance (DISR) and Minimum Redundancy Maximum Relevance (mRMR) can help find the most important features, making classification more accurate and efficient. Together, these findings show how combining different approaches can lead to better and more reliable methods for detecting and classifying ADHD.

While these studies report impressive classification accuracies, the innovation of our proposed method lies in its ability to gain deeper insights into the shared patterns within each group. By focusing on commonalities in EEG signal characteristics across ADHD and typically developing (TD) groups, we aim to uncover the distinctive features that define these populations. This approach has the potential to enhance our understanding of the underlying mechanisms of ADHD and refine classification methods by emphasizing the shared neural patterns unique to each group.

2.4. Introduction to Text Compression

Text compression is a process of reducing the size of a text file or document without losing its essential content. The primary goal is to represent data in a more compact form so that it takes up less space, improves storage efficiency, and allows for faster data transmission.

Apart from decreasing file size, compression of text can also be applied to compare metrics between varying data objects, for instance, text, signals, or sequences. When attempting to measure the compressibility of two objects together, it is possible to define their similarity or any patterns they share. This method will help solve some kinds of problems such as classification, clustering, or pattern recognition since it aids in identifying the relationships within the data. [9].

2.4.1. Formulas

This section introduces two important formulas used in text compression: Normalized Compression Distance (NCD) and Normalized Google Distance (NGD). These formulas are tools for measuring similarity between data, whether it’s text, audio, or other types. NCD uses compression to find how much information is shared between two objects, while NGD measures the similarity of words or phrases based on their frequency in search

results. Both formulas help simplify complex data and find patterns, making them useful in many areas of research and analysis.

- **Normalized Compression Distance (NCD):**

$$\mathrm{NCD}(x,y) = \frac{C(xy) - \min(C(x), C(y))}{\max(C(x), C(y))}$$

Measures similarity between two objects based on their shared redundancy. Lower values indicate higher similarity.

- **Normalized Google Distance (NGD):**

$$\mathrm{NGD}(x,y) = \frac{\max(\log f(x), \log f(y)) - \log f(x,y)}{\log N - \min(\log f(x), \log f(y))}$$

Measures semantic similarity based on co-occurrence of terms in search results.

2.4.2. Definitions

We present definitions that are used when presenting and analyzing articles that use text compression, and maybe we will use some of them in our solution in the future.

1. **Normalized Compression Distance (NCD):-** is a measure that helps identify how similar two objects (like texts or data sequences) are. It does this by using compression algorithms to see how much information two objects share. If two objects are very similar, their NCD will be close to 0. If they are very different, their NCD will be closer to 1.
2. **P300:-** The P300 is a type of brainwave detected in electroencephalography (EEG) that occurs approximately 300 milliseconds after you recognize something important or meaningful. It is part of the brain's natural response to recognizing or paying attention to a specific stimulus, such as a sound, light, or image.

For example, imagine you are participating in an experiment where images flash on a screen, most of them are cars, but occasionally, a picture of a cat appears. When a car flashes, your brain processes it but doesn't find it noteworthy. However, when the cat image appears, your brain recognizes it as important because you're paying attention to it. This recognition triggers a P300 signal about 300 milliseconds after seeing the cat, visible as a positive spike in the EEG.

3. **Support Vector Machines (SVMs):-** are a type of machine learning algorithm used for classification and regression tasks. The main goal of an SVM is to find the best boundary (or hyperplane) that separates data points into different categories.
4. **Normalized Google Distance (NGD):-** is a method for measuring the semantic similarity between words or phrases by analyzing their co-occurrence in Google search results. It is inspired by the Normalized Compression Distance (NCD) but uses search engine data instead of compression.

   The goal: NGD estimates how closely two terms are related by checking how often they appear together in Google search results compared to how often they appear individually. If two terms frequently co-occur in the same search results, they are considered more similar.

5. **Type 1 diabetes (T1D):-** is a condition where the body stops making insulin because the immune system damages the cells that produce it. People with T1D need to take insulin every day to manage their blood sugar levels.
6. **Attributed Tracking Graph (ATG):-** is a way to organize and display information about how objects, like cells or tissues, move and change over time. It shows their interactions and similarities, helping researchers track and understand complex biological processes.
7. **Genetic Algorithm (GA):-** is a method inspired by the process of natural selection and evolution. It is used to solve optimization and search problems by mimicking how biological evolution works.

2.4.3. Articles Using Text Compression

Text compression techniques have shown remarkable potential in identifying patterns and simplifying classification tasks across various domains. One notable study applied Normalized Compression Distance (NCD), a method based on text compression, to detect P300 signals in EEG readings, which are crucial for Brain-Computer Interfaces (BCIs). By converting EEG data into text and using compression algorithms, the study successfully distinguished P300 signals from non-P300 patterns, demonstrating the utility of text compression for improving data representation and advancing BCI technology [15].

Building on the success of NCD, another study explored Normalized Google Distance (NGD) to measure semantic similarity between words or phrases using Google search data. NGD, a feature-free approach, relies on the frequency of term co-occurrences in search results. It has proven effective for grouping items, classifying data with Support Vector Machines (SVMs), translating languages, and validating semantic relationships with WordNet. The study highlighted NGD's robustness, achieving 87.25% accuracy compared to expert classifications, emphasizing its capability to analyze large datasets and understand word relationships [4].

The versatility of compression techniques extends to medical signal analysis, as demonstrated by a study that used Modified Normalized Compression Distance (mNCD) to analyze blood sugar patterns in type 1 diabetes (T1D) patients. By transforming blood sugar data into symbolic representations, mNCD helped identify glycemic control profiles, revealing variations in sugar levels, insulin sensitivity, and hypoglycemia risk. This approach offers a pathway for personalized diabetes management, showing how compression methods can provide actionable insights into complex medical data [6].

Compression-based techniques have also been applied to audio classification. One study used NCD to classify bird species based on audio recordings from the xeno-canto database, an online collection of bird sounds. By bypassing complex feature extraction and directly measuring similarities in audio data, NCD effectively grouped bird species, even in noisy conditions. This underscores the potential of compression methods for analyzing unstructured datasets with minimal preprocessing [14].

In the field of music analysis, researchers applied NCD to group music files by similarity, testing three data representations: WAV format, raw wave information, and Symbolic Aggregate approXimation (SAX). While wave information provided the most accurate results, WAV and SAX had higher errors due to missing or redundant details. The study further introduced a genetic algorithm (GA) to refine the grouping process, the importance of choosing appropriate data representations to enhance classification accuracy [10].

Another innovative application of NCD involved improving text grouping through "annealing text distortion." By replacing common words with symbols, researchers preserved essential information while enhancing grouping accuracy. This approach worked across diverse datasets, showcasing how small adjustments in text can help NCD effectively identify patterns and relationships [11].

Finally, NCD has been used to summarize changes in biological image sequences. By comparing how well data compresses together versus separately, researchers analyzed image sequences to track objects (such as cells or tissues) and their changes over time. The data was organized into an Attributed Tracking Graph (ATG) to visualize interactions and

similarities. This method successfully identified factors like neuroprosthetic insertion speed or neural cell development stages, demonstrating the adaptability of NCD for unsupervised analysis of complex biological processes [5].

Together these studies collectively highlight the versatility and effectiveness of compression-based techniques like NCD, NGD, and mNCD in various domains, including EEG analysis, semantic similarity, medical signals, audio classification, music analysis, text grouping, and biological image summarization. By simplifying complex data and identifying meaningful patterns, these methods provide a universal, parameter-free, and scalable approach for analyzing diverse datasets. Their success across applications underscores the potential of compression-based techniques to drive innovation in data analysis and representation.

2.4.4. Key Findings

The reviewed articles show that text compression techniques can be used in many different fields to make data easier to analyze and understand. In EEG studies, methods like Normalized Compression Distance (NCD) have been used to detect P300 signals, helping improve Brain-Computer Interfaces (BCIs) and the study of brain patterns. Normalized Google Distance (NGD) has been used to measure the relationship between words or phrases, achieving high accuracy in tasks like grouping, classification, and translation. In medical research, Modified NCD (mNCD) has helped analyze blood sugar patterns in diabetes patients, offering better ways to manage their health. NCD has also been used to classify bird sounds and group music files, showing it can handle complex data without extra processing. Other creative uses, like improving text grouping and summarizing biological images, show how compression techniques can reveal patterns and relationships in data. Overall, these studies highlight how useful and flexible text compression methods can be for solving problems in many areas.

3. Our solution

Our solution involves classifying EEG recordings by calculating similarity scores between signals from the same channel and analyzing these scores to determine group membership (ADHD or typically developing).

3.1. Preprocessing

The EEG dataset consists of raw recordings that require preprocessing to ensure data quality before being used as input for our algorithm. This preprocessing step is crucial to remove noise and artifacts, and prepare it for analysis. By working directly with the raw data, we aim to maintain the integrity of the original signals and maximize the algorithm's effectiveness. The preprocessing includes:

Noise Reduction: Filtering out artifacts and unwanted noise from the signals while retaining essential information.

3.2. Similarity Matrix Construction

For each EEG channel, we construct a 120x120 similarity matrix that represents the similarity scores between every pair of signals in the dataset. The process is as follows :

1. **Iterate Through Channels**:

For each of the 19 EEG channels (e.g., FP1, FP2, Fz, etc.), perform the following steps:

1.1. **Iterate Through Signals**:

For each signal in the channel (from 1 to 120), calculate its similarity with all other signals in the same channel.

1.1.1. **Calculate Similarity Scores**:

For each pair of signals, compute a similarity score using a text compression algorithm, as you can see in Fig. 6. Store the result in the corresponding cell of the similarity matrix.

The step-by-step process of constructing the similarity matrix is outlined in the pseudo-code below *.

This results in 19 similarity matrices (one for each channel), each capturing the relationships between signals for that specific channel. Fig. 7. presents an example of a 120x120 similarity matrix for one EEG channel, it is not final yet for how it looks.

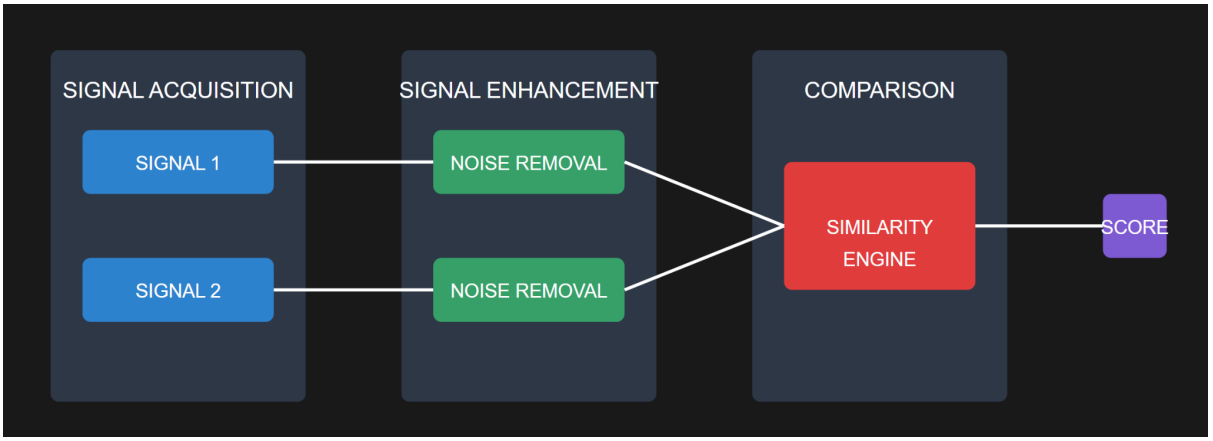

Fig. 6. This figure explains the process of our solution. First of all, the two signals undergo a noise removal process, then they enter our algorithm. The score of the output defines whether signal 2 is matched or unmatched for signal 1.

3.3. Classification

After building the similarity matrices, we analyze the stored similarity scores to classify participants into their groups (ADHD or typically developing). The steps are as follows:

1. **Analyze Patterns**:
   - Examine the similarity scores within each matrix to identify patterns that differentiate the ADHD group from the typically developing group.
2. **Group Classification**:
   - Use the similarity scores to classify each signal into one of the two groups.
   - Repeat the analysis across all 19 channels to refine the classification accuracy.
3. **Aggregate Results**:
   - Combine the classification results from all channels to make a final determination about the participant's group membership.

In the classification step, the similarity scores from the matrices are analyzed as follows: For each signal, we examine its 120 similarity scores from the corresponding similarity matrix. If the signal belongs to the ADHD group, the 60 highest similarity scores are classified as ADHD, while the 60 lowest similarity scores are classified as typically developing. Conversely, if the signal belongs to the typically developing group, the 60 highest similarity scores are classified as typically developing, and the 60 lowest scores are classified as ADHD.

*** The process can be represented as pseudo-code:**

```
for channel in 1 to 19:
    for signal_1 in 1 to 120:
        for signal_2 in 1 to 120:
            similarity_score = calculate_similarity(signal_1, signal_2)
            store_similarity_score(channel, signal_1, signal_2, similarity_score)
```

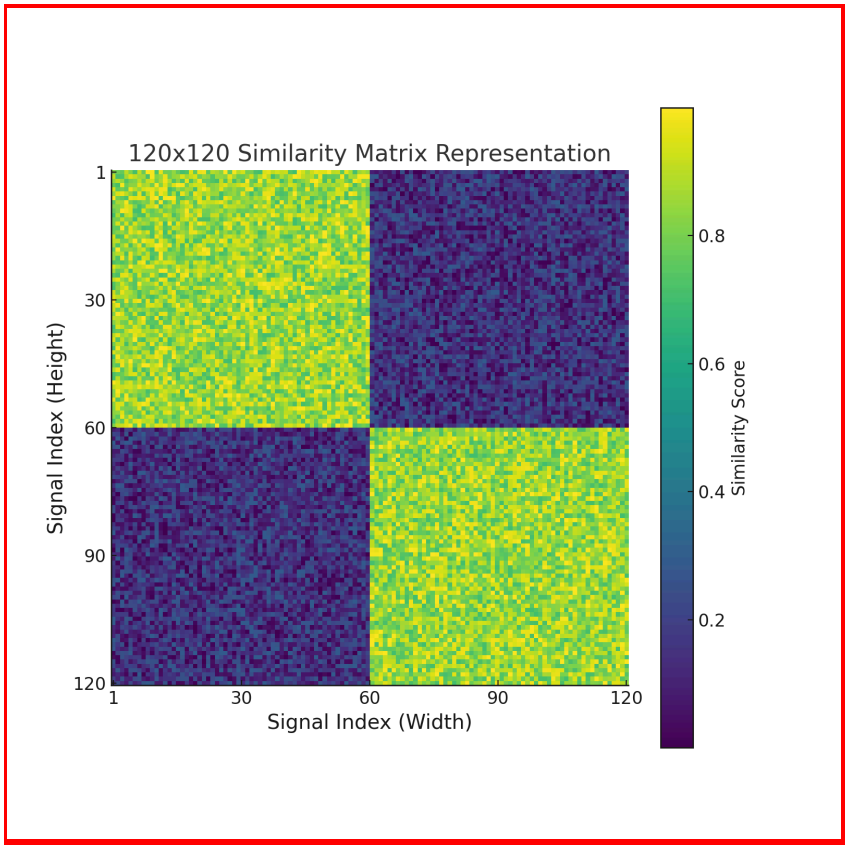

Fig. 7. The 120x120 similarity matrix shows how similar each signal is to every other signal. The large yellow blocks mean the signals in those areas belong to the same group and are very similar. The color scale on the right shows that dark blue means low similarity and bright yellow means high similarity.

3.4. Testing Methodology and Evaluation

To ensure our EEG classification system works correctly, we set several testing goals and follow clear steps to evaluate them. First, we check the accuracy of our classification algorithm by comparing its results to a labeled dataset where the ADHD and typically developing signals are already known.

Second, we test whether the preprocessing step successfully removes noise while keeping important signal details. We do this by comparing EEG signals before and after preprocessing to confirm that unwanted noise is reduced without losing useful information.

Third, we verify that the similarity matrices are correct by manually calculating similarity scores for a few signals and checking if they match the computed results. We also use visual analysis to ensure that signals from the same group have higher similarity scores than those from different groups.

Finally, we test how well the system handles challenges like missing data, noisy signals, and extreme values. This helps us confirm that our method remains reliable even when dealing with imperfect data. These tests ensure our system is accurate, effective, and ready for real-world use.

3.5. Anticipated Outcomes

Our solution is expected to accurately classify EEG recordings into ADHD and typically developing groups by identifying unique patterns in brain activity. Using similarity matrices built from raw EEG data with text compression algorithms, we can preserve and analyze detailed information that might be missed with traditional methods.

This approach will not only provide high classification accuracy but also reveal important differences in how the brain works in children with ADHD compared to those without it. These findings can improve our understanding of ADHD-related brain activity and could support the development of better diagnostic tools and treatment monitoring systems.

By analyzing data from all 19 EEG channels and combining the results, our method ensures a comprehensive evaluation of brain activity. This technique has the potential to reduce misdiagnoses and enhance personalized treatment plans for ADHD. Additionally, our solution highlights how text compression can be used in neuroscience to analyze complex data, setting a foundation for its application in other areas, such as medical diagnosis, cognitive research, and brain-computer interfaces.

AI Tools Used

We used AI tools to support various aspects of our research and planning. ChatGPT was instrumental in generating creative and innovative ideas, helping us explore different approaches and refine the direction of our project. It provided valuable insights that shaped our methodology and research focus. ChatPDF assisted us in analyzing academic articles, allowing us to extract key information efficiently and understand relevant research more effectively. By using these AI tools, we improved our research efficiency and gained a clearer understanding of how to implement our solution.

References

[1] Adamou, M., Fullen, T., & Jones, S. L. (2020). EEG for diagnosis of adult ADHD: a systematic review with narrative analysis. *Frontiers in Psychiatry, 11,* 871.

[2] Allahverdy, A., Moghadam, A. K., Mohammadi, M. R., & Nasrabadi, A. M. (2016). Detecting ADHD children using the attention continuity as nonlinear feature of EEG. *Frontiers in Biomedical Technologies, 3(1-2),* 28-33.

[3] Biasiucci, A., Franceschiello, B., & Murray, M. M. (2019). Electroencephalography. *Current Biology, 29(3),* R80-R85.

[4] Cilibrasi, R. L., & Vitanyi, P. M. (2007). The Google Similarity Distance. *IEEE Transactions on Knowledge and Data Engineering, 19(3),* 370-383.

[5] Cohen, A. R., Bjornsson, C. S., Temple, S., Banker, G., & Roysam, B. (2008). Automatic summarization of changes in biological image sequences using algorithmic information theory. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(8),* 1386-1403.

[6] Contreras, I., Quirós, C., Giménez, M., Conget, I., & Vehi, J. (2016). Profiling intra-patient type I diabetes behaviors. *Computer Methods and Programs in Biomedicine, 136,* 131-141.

[7] Da Silva, B. S., Grevet, E. H., Silva, L. C. F., Ramos, J. K. N., Rovaris, D. L., & Bau, C. H. D. (2023). An overview on neurobiology and therapeutics of attention-deficit/hyperactivity disorder. *Discover Mental Health, 3(1),* 2.

[8] Ekhlasi, A., Nasrabadi, A. M., & Mohammadi, M. R. (2021). Direction of information flow between brain regions in ADHD and healthy children based on EEG by using directed phase transfer entropy. *Cognitive Neurodynamics, 15(6),* 975-986.

[9] Fitriya, L. A., Purboyo, T. W., & Prasasti, A. L. (2017). A review of data compression techniques. *International Journal of Applied Engineering Research, 12(19),* 8956-8963.

[10] González-Pardo, A., Granados, A., Camacho, D., & de Borja Rodríguez, F. (2010, July). Influence of music representation on compression-based clustering. In *IEEE Congress on Evolutionary Computation (pp. 1-8). IEEE.*

[11] Granados, A., Cebrian, M., Camacho, D., & de Borja Rodriguez, F. (2010). Reducing the loss of information through annealing text distortion. *IEEE Transactions on Knowledge and Data Engineering, 23(7),* 1090-1102.

[12] Khazi, M., Kumar, A., & Vidya, M. J. (2012). Analysis of EEG using 10: 20 electrode system. *International Journal of Innovative Research in Science, Engineering and Technology, 1(2),* 185-191.

[13] Mohammadi, M. R., Khaleghi, A., Nasrabadi, A. M., Rafieivand, S., Begol, M., & Zarafshan, H. (2016). EEG classification of ADHD and normal children using non-linear features and neural network. *Biomedical Engineering Letters, 6,* 66-73.

[14] Sarasa, G., Granados, A., & Rodriguez, F. B. (2017, September). An approach of algorithmic clustering based on string compression to identify bird songs species in xeno-canto database. In *2017 3rd International Conference on Frontiers of Signal Processing (ICFSP) (pp. 101-104). IEEE.*

[15] Sarasa, G., Granados, A., & Rodriguez, F. B. (2019). Algorithmic clustering based on string compression to extract P300 structure in EEG signals. *Computer Methods and Programs in Biomedicine, 176,* 225-235.

[16] Talebi, N., & Nasrabadi, A. M. (2022). Investigating the discrimination of linear and nonlinear effective connectivity patterns of EEG signals in children with Attention-Deficit/Hyperactivity Disorder and Typically Developing children. *Computers in Biology and Medicine, 148,* 105791.

[17] Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review, 2(2),* 1-11.

[18] https://www.cdc.gov/adhd/diagnosis/index.html?utm_source

[19] https://ieee-dataport.org/open-access/eeg-data-adhd-control-children