Neurodivergent AI Assistant

Vinoth Arumugam
220676058
Prof. Ioannis (Yiannis) Patras
MSc Artificial Intelligence

ABSTRACT

Identifying autism spectrum disorder (ASD) in children, a common mental condition, is crucial in medical care. Recently, graph neural networks (GNN) based on functional brain networks (FBN) have been effective in disease diagnosis. Yet, challenges in creating the best FBN from resting-state fMRI data remain. Also, it's unclear how different FBN structures affect GNN-based disease classification. Our study uses Dynamic Time Warping (DTW) to measure FBN connectivity. These are turned into brain graphs and then analysed by a graph attention network (GCN) to spot ASD. We tested various FBN methods and classification models on the ABIDE dataset (n = 1112). The results highlighted the effectiveness of the DTW method and how different FBNs impact GNN-based results. Using DTW and GCN, we achieved an ASD classification accuracy of 60%, close to current methods. This progress will help differentiate patients from controls, showing promise for future disease detection using FBN and GNN on bigger, more complex datasets.

Keyword: Functional Brain Network(FBN), Autism Brain Image Data Exchange(ABIDE), Graph Convolutional Neural Netwrok(GCN),
Functional Magnetic Resonance Image(FMRI), Blood Oxygen Level Dependent(BOLD), Autism Spectrum Disorder(ASD), Configurable Pipeline for the Analysis of Connectomes(CPAC)

I. Introduction:

Autism Spectrum Disorder (ASD) is a neurological developmental issue in children characterized by difficulties in social interaction, language use, limited interests, and repetitive actions [1]. At present, diagnosing ASD mainly depends on visible symptoms and the assessment of medical experts, which can result in incorrect diagnoses. Hence, there's an essential demand for an accurate automated system for ASD diagnosis. Thanks to modern magnetic resonance imaging (MRI) techniques, generating a functional brain network (FBN) using resting-state functional MRI (rs-fMRI) provides a reliable method to identify minor unusual shifts in brain conditions. FBN shows the functional links between different brain areas, based on the relationship of time sequences in those regions. This method can accurately identify specific active brain zones and track signal alterations instantly [2–4]. Correctly establishing this brain functional network is crucial for subsequent statistical analysis and disease classification. Several research confirm that FBN analysis can reveal unique healthy brain structures and highlight key markers to differentiate between neurological and psychological disorders, including ASD [5,6].

II. Background:

The Autism Brain Imaging Data Exchange (ABIDE) is a significant initiative in the field of neuroscience. It was established to promote collaboration among researchers by providing a platform where they can share functional magnetic resonance imaging (fMRI) datasets related to autism spectrum disorder (ASD). By pooling data from various global sources, ABIDE aims to facilitate a deeper understanding of ASD, allowing for more comprehensive analyses than individual studies might offer.

The Configurable Pipeline for the Analysis of Connectomes (CPAC) is closely associated with the ABIDE initiative. CPAC is a software tool designed to simplify the processing and analysis of fMRI data. Given the complexity and variety of fMRI datasets, having a standardized pipeline like CPAC is crucial. It ensures that data from different sources can be processed in a consistent manner, making comparisons and aggregations more valid. In essence, CPAC provides researchers with a set of tools and workflows to analyze brain connectivity, making it an invaluable resource for those working with ABIDE datasets and beyond.

Brain Network Methodologies and Their Importance:
Studies have shown that a well-designed brain network(FBN) approach can improve the precision of subsequent disease identification[8]. The term FBN refers to the Functional Brain Network. This network represents the functional connections or interactions between different brain regions. By studying the FBN, researchers can understand how different brain regions communicate and work together, especially in individuals with ASD. The FBN is crucial as abnormalities or differences in these networks can provide clues about the neurological underpinnings of autism.

Pearson's correlation (PC) is the favoured technique for establishing FBN, capturing all links between

brain regions [8]. While PC's calculation is direct and biologically meaningful, it doesn't account for potential interference from other brain areas, possibly leading to incorrect links.

Challenges and Solutions in Brain Network Construction:

Partial correlation addresses interference by accounting for potential confounders using regression. However, this method can be problematic, especially when there are fewer time points than brain regions. To achieve consistent outcomes, regularization methods like sparse representation (SR) [9] are often used. SR can also reduce weak connections due to noise without setting arbitrary thresholds. Lately, researchers have proposed adding regularization terms to the SR model to include prior knowledge of the brain network. Examples include the work of Lee et al., Varoquaux et al., Wee et al., Qiao et al., and Xue et al. [10-15]. However, some research suggests that SR-based brain networks might be less effective than PC-based ones [6,16]. Additionally, the PC method avoids certain challenges present in the SR method, offering greater statistical robustness and scalability.

Classification Methods in Disease Diagnosis:

A critical factor in disease diagnosis using FBN is the classification method chosen. Existing studies often use conventional machine learning techniques like random forest (RF) and support vector machine (SVM) [13,15,17], which might restrict their potential. This limitation stems from issues such as:
1. The difficulty of using traditional machine learning for large datasets.
2. The intricate process of feature engineering in machine learning which involves initial data exploration, dimension reduction, and optimal feature selection.

Deep Learning's Role in Brain Network Analysis:

In contrast to traditional methods, deep learning models are advantageous as they eliminate the need for intricate feature engineering, directly inputting data and often producing superior results. They can also be easily tailored to various applications. Another challenge is the problem of increasing data dimensions as tasks become more complex, especially with data having numerous features, traditional machine learning struggles. Recently, the spatial-temporal characteristics of fMRI have been used to depict the brain's fMRI in a graph structure, preserving both space and time elements.

Graph Neural Networks in Disease Detection:

Graph neural network (GNN), a fitting deep learning method for graphs, has shown significant success in disease detection [18–20]. Examples include the work of Ktena et al., Jiang et al., Yao et al., and Li et al. [21-23]. These studies highlight GNN's advantage over traditional machine learning in analyzing brain networks. Yet, how different brain network methods influence GNN-based disease detection remains unexplored.

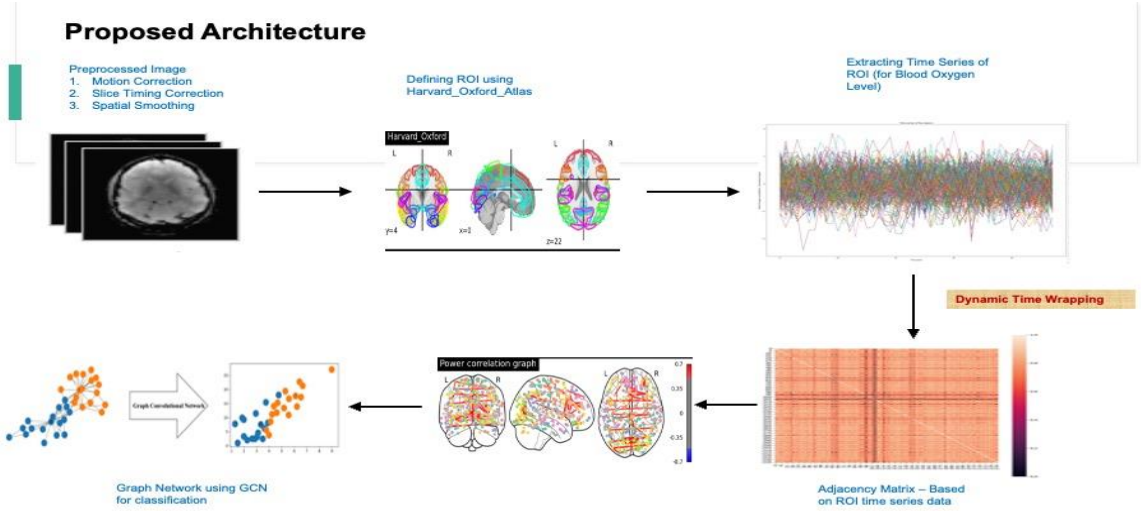
This research implemented a technique for creating FBN called Dynamic Time Wrapping (DTP) to shape the data and incorporate knowledge effectively. Several tests were carried out on the graph neural network's graph structure, revealing that different brain network creation techniques can impact the success of GNN-based disease categorization. All these tests utilized fMRI data from the Autism Brain Imaging Data Exchange (ABIDE) I dataset, and the results from the ABIDE dataset validate the performance of the suggested DTW and GCN classification model.

Regarding Functional Brain Networks (FBN), the assessment typically comes from the correlation between Blood Oxygen Level Dependent (BOLD) signals in different brain areas. Methods like Pearson Correlation (PC), partial correlation, and sparse representation (SR) have been employed for this task, along with some intricate higher-order variables [25]. However, since correlation-focused methods have proven effective in real-world scenarios and have shown more sensitivity than the advanced higher-order techniques [16], this report will focus exclusively on the DTW approach.

The PC method constructs a functional connectivity network by measuring the complete correlation among various Regions of Interest (ROIs). Its widespread use is due to its straightforwardness, statistical strength, and computational speed, making it the preferred choice for many. However, it's limited to capturing a basic linear relationship and doesn't consider the biological processes behind brain disorders [17].

For the sake of clarity and without loss of generality, we assume that the mean time series of *ith* ROI is x_i, and the series of the whole brain is denoted as $X=\{x1,x2,\cdots\cdots,xN\} \in R^{T\times N}$ (N is the number of brain regions and T is the number of time points in each mean time series), $W \in R^{N\times N}$ is the estimated strength of connections between brain regions in all of the following equations. PC can be defined as:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$



This framework laid the foundation on the research community to perform various analyses on the brain's functional connectivity, helping understand normal brain function and various neurological and psychiatric disorders. The simplicity of the PC method, in particular, allows for robust and efficient computations, although its linear modelling may be seen as a limitation when seeking to capture more complex relationships within the brain.

Partial Correlation as an Alternative to PC,
the Partial Correlation -based scheme is used as a standard of comparison against our newly suggested technique. The mathematical model for Partial Correlation can be described by a particular objective function [10].

$$r_{xy.z} = \frac{r_{xy} - r_{xz} \cdot r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}}$$

III. Methodology (Figure 1):

Data Acquisition and Preprocessing:

The foundation of this study was laid with the acquisition of the ABIDE dataset, which underwent processing via the CPAC pipeline. This dataset combines 1112 functional MRI (fMRI) images from different sites, supplemented by phenotypic data presented in a CSV format. The dataset was divided into training and testing subsets to ensure a systematic approach to model training and evaluation.

MRI Processing and Parcellation:

Once the data is available the next step involved the processing of fMRI images. Each subject's MRI underwent a label masking procedure, facilitated by the Harvard Oxford atlas. This atlas played a pivotal role in the parcellation of the brain, enabling the identification and differentiation of distinct brain regions.

Time Series Extraction and Connectivity Analysis:

With the brain regions clearly identified, the subsequent phase was dedicated to the extraction of time series based on Blood Oxygen Level Dependent (BOLD) signals. These BOLD signals are instrumental in gauging the brain's activity during its resting states across varied time frames. To delve deeper into the relationships between different brain regions, Dynamic Time Warping (DTW) was employed, constructing a connectivity measure illustrate the complicated relationship between these regions.

Graph Construction:

The connectivity metrics served as the blueprint for the construction of a directed graph. In this graphical representation, the nodes epitomized
the brain regions, while the edges, with their
varying weights, depicted the strength of
connectivity between these regions.

Neural network analysis:

The final phase of the methodology was anchored in the realm of neural network analysis. The previously constructed graph was fed into a Graph Neural Network, with the GCN (Graph Convolutional Network) being the primary network underpinning this study. The overarching aim of this neural network analysis was the classification of subjects, discerning between autistic individuals and controlled counterparts.

This structured methodology provides a comprehensive roadmap, guiding the study from

data acquisition to neural network analysis, all with the aim of collecting insights into autism from the ABIDE dataset. The study embarks on an elaborate analysis of MRI data sourced from the Autism Brain Imaging Data Exchange (ABIDE) CPAC pipeline, employing a composite pipeline that integrates state-of-the-art machine learning and graph theory techniques.

IV. Implementation

Design:

To commence the implementation, it's essential to integrate the necessary functionalities from the **nilearn.datasets** module. The function **fetch_abide_pcp** emerges as the prerequisite in this context. Its role is to facilitate the retrieval of data from the esteemed ABIDE (Autism Brain Imaging Data Exchange) collection.

By setting the pipeline parameter to **"cpac"**, the function is oriented to source data that has undergone the rigorous preprocessing steps characteristic of CPAC, ensuring data consistency and reliability.

Basic Processing:

a. **Slice Timing Correction**:
Adjusts the timing differences in MRI slices to make them consistent.

b. **Motion Realignment:**
Corrects small movements that happen during the MRI scan.

c. **Intensity Normalization:**
Standardizes the brightness levels across MRI images.

Nuisance Signal Removal:

a. **24-Parameter Regression:**
Removes unwanted motion-related signals.

b. **Tissue Signals (CompCor):**
Reduces noise from certain brain areas.

c. **Motion Realignment:**
Further corrects for any movement during the scan.

d. **Low-Frequency Drifts:**
Removes slow, unrelated signal changes in the MRI data.

Processing Strategies:

a. **Band-Pass Filtering:**
Keeps only the important brain signals and removes unwanted frequencies.

b. **Global Signal Regression**:
Reduces general noise to focus on specific brain signals.

Registration:

a. **Functional to Anatomical:**
Aligns brain activity images with individual's brain structure.

b. **Anatomical to Standard:**
Matches the individual's brain images to a common template for easier comparison.

In essence, the CPAC pipeline fine-tunes MRI data, making it clearer and more suitable for analysis.

The integrity of the data is vital. By enabling the **quality_checked** parameter **(set to True),** the function is directed to prioritize datasets that have been vetted and have passed stringent quality assessments. This is a crucial step, ensuring that the foundation of subsequent analyses is robust and dependable.

Then the MRI data is loaded into the environment using the **nib.load** function. This function reads the MRI image, making its details accessible for subsequent operations. The loaded image is stored in the variable.

To further the neuroimaging processing, the **datasets.fetch_atlas_harvard_oxford** function is employed. The Harvard-Oxford (HO) atlas, which comes with FSL, is divided into probabilistic atlases for both the cortex and subcortex. Both atlases had a 25% threshold applied. They were then split into left and right hemispheres at the central line (x=0). From the subcortical atlas, regions representing left/right white matter (WM), left/right gray matter (GM), left/right cerebrospinal fluid (CSF), and the brainstem were excluded. After combining the cortical and subcortical regions of interest (ROIs), they were resized to match the functional resolution using a method called nearest-neighbor interpolation.

This function retrieves the specified Harvard-Oxford atlas, storing it in the variable. With the atlas in hand, the next step is to create a masker using the **input_data.NiftiLabelsMasker** function. A vital part of this phase is the application of a `NiftiLabelsMasker` to the MRI images, a step that calls for the utilization of the atlas_harvard_oxford to parcellate the brain into discrete regions. This meticulous parcellation is instrumental in defining the structural and functional details of the brain,

thereby laying a solid foundation for the subsequent analytical stages. The Nifti data can then be turned to time-series by calling the `**NiftiLabelsMasker.fit_transform**` method, that takes either filenames or **NiftiImage objects.**

The important technique lies in the construction of a distance matrix. This matrix aims to capture the distances or dissimilarities between time series data of different brain regions. The **fastdtw** function, a method for **Dynamic Time Warping(DTW),** is employed to compute these distances. By iterating over all pairs of regions, the distance between their respective time series data is computed and stored in the distance_matrix.

DTW is a renowned time-series alignment technique, and in this context, it's wielded to gauge the similarity between the time series of different ROIs. The mathematical underpinning of DTW can be encapsulated in the following recursive formula:

$$D(i,j) = d(i,j) + \min\{D(i-1,j), D(i,j-1), D(i-1,j-1)\}$$

Here, the accumulated distance D(i,j) is computed as the sum of the local distance d(i,j) and the minimum of the previous accumulated distances, considering three possible moves:

1. Horizontal Move: This move aligns the current element in the first sequence with a gap in the second sequence, represented by the accumulated distance $(D(i-1,j))$.

2. Vertical Move: This move aligns the current element in the second sequence with a gap in the first sequence, represented by the accumulated distance $(D(i,j-1))$.

3. Diagonal Move: This move aligns the current elements in both sequences with each other, represented by the accumulated distance $(D(i-1,j-1))$.

These three options represent all the possible ways to proceed in aligning the sequences at the current position, and the algorithm selects the option that minimizes the accumulated distance.

Foundation of the Model:

The `Net` class is derived from `torch.nn.Module`, a base class provided by PyTorch for all neural network modules. This inheritance allows for the seamless integration of layers, functions, and methods that PyTorch offers, ensuring that the model can be trained, evaluated, and optimized efficiently.

Layered Architecture (Figure 2):

Net(
  (conv1): GCNConv(48, 16)
  (bn1): BatchNorm1d(16, eps=1e-05, momentum=
0.1, affine=True, track_running_stats=True)
  (conv2): GCNConv(16, 32)
  (bn2): BatchNorm1d(32, eps=1e-05, momentum=
0.1, affine=True, track_running_stats=True)
  (fc): Linear(in_features=32, out_features=2, bias=
True)
  (dropout): Dropout(p=0.5, inplace=False)
)

Graph Convolutional Network (GCN) Layers: The model incorporates two GCN layers, denoted as `conv1` and `conv2`. The GCN is a type of neural network layer

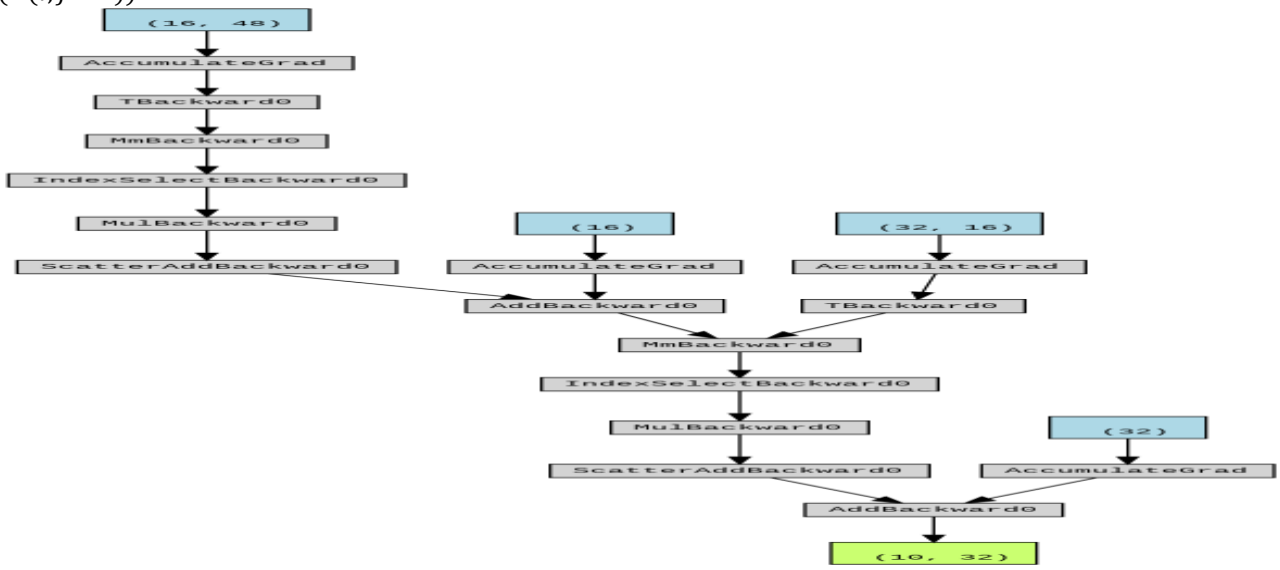


Figure 2

specifically designed to operate over nodes in a graph.
The first layer, `conv1`, takes the node features as input and transforms them into a 16-dimensional space. This transformation allows the model to capture and learn more complex patterns from the input features.
The second layer, `conv2`, further processes the 16-dimensional data from the previous layer into a 32-dimensional space. This hierarchical processing ensures that the model can capture multi-scale patterns in the data.

Batch Normalization:
Following each GCN layer are batch normalization layers (`bn1` and `bn2`). Batch normalization is a technique that normalizes the activations of the nodes, ensuring that they have a consistent mean and variance. This normalization accelerates training and provides some regularization, reducing the risk of overfitting.

Dropout for Regularization:

The `dropout` layer is introduced to prevent the model from becoming too reliant on any specific node or feature. By randomly setting a fraction (in this case, 50%) of the input units to 0 at each update during training, dropout helps in preventing overfitting. This ensures that the model generalizes well to new, unseen data.

Pooling:

The **global_mean_pool** function is applied, which aggregates the features of all nodes in a graph into a single global feature vector. This step reduces the dimensionality of the data, making it suitable for classification.

Final Linear Layer:

The `fc` layer is a fully connected (or linear) layer that transforms the 32-dimensional data from the previous layers
into a space that matches the number of classes (e.g., autistic or non-autistic). This layer essentially decides the final classification based on the patterns and features recognized by the preceding layers.

Model Parameters:

Batch Size – 32
Epoch – 100
Learning rate - [0.01, 0.001, 0.0001]
Optimizer – Adam
Regularization – L2('5e-4')
Loss Function - negative log likelihood loss **(F.nll_loss)**

Result:

Connectivity Measure	ACC	PRE	REC	F1	SUP
Pearson Correlation (0.01)	0.63	0.64	0.63	0.60	186
Partial Correlation (0.001)	0.60	0.61	0.60	0.60	186
DTW(0.0001)	0.58	0.58	0.58	0.58	186

Table:1

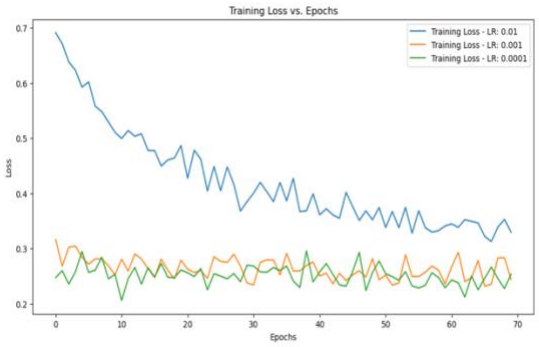


Figure 3 – DTW Technique

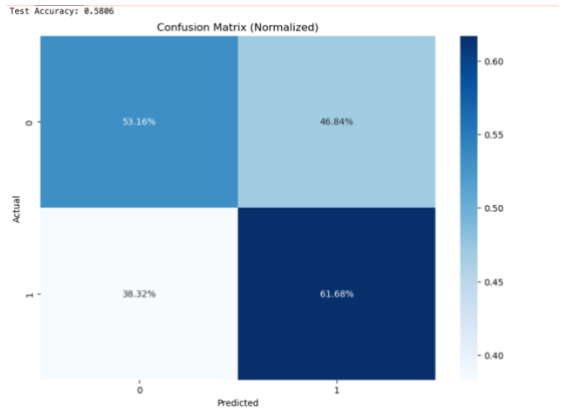


Figure 4 – DTW Technique

VII. Critical Analysis:

Pearson Correlation is a straightforward measure that quantifies the strength and direction of a linear relationship between two variables. Its simplicity and interpretability are its main advantages. However, it only captures linear relationships and can be sensitive to outliers. It's best suited for stationary time series data. Given its widespread use, it's often the first choice in many statistical analyses.

Partial Correlation, on the other hand, offers a more nuanced view by controlling for confounding variables. This allows for a clearer understanding of relationships by removing the effects of other variables. While it provides a more detailed

perspective than Pearson, it's also more complex and assumes linearity among variables. Like Pearson, it's most applicable to stationary time series.

Dynamic Time Warping (DTW) is a versatile measure, especially adept at comparing sequences of different lengths and timings. It can capture non-linear relationships, making it more flexible than the other two measures. However, its computational demands can be high for long sequences and might produce boundary artefacts. DTW's ability to handle both stationary and non-stationary time series makes it a preferred choice for diverse time series analyses.

In this study, we explored the efficacy of Dynamic Time Warping (DTW) compared to baseline connectivity measures. Despite the theoretical advantages of DTW, particularly its suitability for time series data, the results did not significantly improve over the baseline. Several factors could account for this observation:

Data Limitations:
Insufficiency: The dataset, comprising only 48 data points for each of the 240+ regions, may not adequately represent the intricate dynamics or relationships between these regions. Such a limited dataset can compromise the robustness of models, leading to suboptimal performance on novel data.

Dimensional Challenges:

The inclusion of over 240 regions translates to high dimensionality. This scenario can invoke the "curse of dimensionality", where data sparsity becomes a concern due to the vastness of the space. Such sparsity can hinder methods that rely on statistical significance.

Risk of Overfitting:

The disproportionate ratio of regions to data points heightens the risk of overfitting. Consequently, models might become overly attuned to training data nuances, impairing their generalisation capabilities on new datasets.

While our dataset's 48 points were confirmed to be stationary using the Augmented Dickey–Fuller test (ADF), technically, just 48-time points are insufficient to validate ADF, but DTW's performance remained on par with baseline measures. This is intriguing, given DTW's inherent design to excel with time series data. One might argue in favour of baseline techniques due to their computational efficiency and equivalent performance. However, the emphasis isn't solely on accuracy in critical domains like medicine. The rationale behind conclusions is equally paramount.

VIII. Conclusion:

Given this, even if baseline techniques offer faster computation and similar accuracy, their lack of a logical foundation makes them less compelling. In contrast, DTW, despite its current performance, provides a more reasoned approach suitable for medical interpretations. Thus, this study recommends adopting DTW over the previously employed baseline methods for future endeavours.

IX. Future Work:

Granger Causality Test

Explanation:
The Granger causality test is a statistical hypothesis test used to determine whether one time series can predict another time series. It is based on the principle that if variable \(X\) Granger-causes variable \(Y\), then past values of \(X\) should contain information that helps predict \(Y\). It's important to note that the term "Granger causality" does not imply true causality in the traditional sense, but rather a predictive capability.

Suitability for Your Use Case:
1. Time Series Analysis: If you are dealing with time series data, the Granger causality test can be a valuable tool to understand the relationships between different series.

2. Predictive Insights: The test can provide insights into which variables (or time series) might be useful predictors for other variables. This can be particularly useful in multivariate time series analysis.

3. Lag Analysis: The test considers various lags of the time series to determine if one series is predictive of another. This can help in understanding the delayed effect of one variable on another.

Drawbacks:

1. Predictive, Not Causal: The name "Granger causality" can be misleading. A significant result does not imply a true causal relationship in the traditional sense, but rather that one series can predict another.

2. Stationarity Requirement: The time series data must be stationary for the Granger causality test to be valid. This means that the properties of the series (like mean and variance) do not change over time. If the series is non-stationary, it needs to be transformed before applying the test.

3. Multivariate Complexity: While the test can be extended to multivariate scenarios, it becomes more

complex and computationally intensive as the number of variables increases.

4. Lag Selection: The choice of the maximum lag to consider can influence the results. An inappropriate choice might lead to misleading conclusions.

In conclusion, while the Granger causality test can provide valuable insights into the predictive relationships between time series, it's essential to be aware of its limitations. Ensure that the data meets the test's assumptions, and interpret the results with caution, especially when considering real-world implications or interventions based on the findings.

References:

[1] U. Frith, F. Happ´e, Autism spectrum disorder, Curr. Biol. 15 (2005) 786–790.

[2] E.A. Allen, E. Damaraju, S.M. Plis, E.B. Erhardt, Eichele Tom, V.D. Calhoun, Tracking whole-brain connectivity dynamics in the resting state, Cerebr. Cortex 24 (2014) 663–676.

[3] R.A. Poldrack, Region of interest analysis for fMRI, Soc. Cognit. Affect Neurosci. 2 (2007) 67–70.

[4] N.K. Logothetis, What we can do and what we cannot do with fMRI, Nature 453 (2008) 869–878.

[5] Yu Zhang, Han Zhang, Xiaobo Chen, Mingxia Liu, Xiaofeng Zhu, Seong-Whan Lee, Dinggang Shen, Strength and similarity guided group-level brain functional network construction for MCI diagnosis, Pattern Recogn. 88 (2019) 421–430.

[6] Weikai Li, Zhengxia Wang, Limei Zhang, Lishan Qiao, Dinggang Shen, Remodeling pearson's correlation for functional brain network estimation and autism spectrum disorder identification, Front. Neuroinf. 11 (2017).

[7] S.M. Smith, D. Vidaurre, C.F. Beckmann, Matthew F. Glasser, Mark Jenkinson, Karla L. Miller, Thomas E Nichols, Emma C. Robinson, Gholamreza Salimi-Khorshidi, Mark W. Woolrich, Deanna M Barch, Kamil Ugurbil, David C. Van Essen, Functional connectomics from resting-state fMRI, Trends Cognit. Sci. 17 (2013) 666–682.

[8] S.M. Smith, K.L. Miller, G. Salimi-Khorshidi, et al., Network modelling methods for FMRI, Neuroimage 54 (2011) 875–891.

[9] Shuai Huang, Li Jing, Liang Sun, Teresa Wu, Kewei Chen, Fleisher Adam, M. Eric, Reiman and jieping Ye. Learning brain connectivity of alzheimer's disease from neuroimaging data, Adv. Neural Inf. Process. Syst. 22 (2009) 808–816.

[10] H. Lee, D.S. Lee, H. Kang, B. Kin, M.K. Chung, Sparse brain network recovery under compressed sensing, IEEE Trans. Med. Imag. 30 (2011) 1154–1165.

[11] C.Y. Wee, P.T. Yap, D. Zhang, Daoqiang Zhang, Lihong Wang, Dinggang Shen, Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification, Brain Struct. Funct. 219 (2014) 641–656.

[12] G. Varoquaux, A. Gramfort, J.B. Poline, et al., Brain covariance selection: better individual functional connectivity models using population prior, Adv. Neural Inf. Process. Syst. 23 (2010).

[13] Lishan Qiao, Han Zhang, Minjeong Kim, Shenghua Teng, Limei Zhang, Dinggang Shen, Estimating functional brain networks by incorporating a modularity prior, Neuroimage 141 (2016) 399–407.

[14] E. Bullmore, O. Sporns, The economy of brain network organization, Nat. Rev. Neurosci. 13 (2012) 336–349.

[15] Yanfang Xue, Limei Zhang, Lishan Qiao, Dinggang Shen, Estimating sparse functional brain networks with spatial constraints for MCI identification, PLoS One 15 (2020).

[16] Fanglin Huang, Ee-Leng Tan, Peng Yang, Shan Huang, Le Ou-Yang, Jiuwen Cao, Tianfu Wang, Baiying Lei, Self-weighted adaptive structure learning for ASD diagnosis via multi-template multi-center representation, Med. Image Anal. (2020) 101662.

[17] B. Xia-An, Yang Wang, Qing Shu, Qi Sun, Qian Xu, Classification of autism spectrum disorder using random support vector machine cluster, Front. Genet. 9 (2018).

[18] Xiaoxiao Li, Nicha C. Dvornek, Yuan Zhou, Juntang Zhuang, Pamela Ventola, James S. Duncan, Graph Neural Network for Interpreting Task-fMRI Biomarkers, 11768, 2019, pp. 485–493.

[19] Xiaoxiao Li, Yuan Zhou, Siyuan Gao, Nicha Dvornek, Muhan Zhang, Juntang Zhuang, Shi Gu, Dustin Scheinost, Staib Lawrence, Pamela Ventola, James Duncan, BrainGNN: interpretable brain graph neural network for fMRI Analysis, Med. Image Anal. (2021) 102233.

[20] Hao Jiang, Peng Cao, MingYi Xu, Jinzhu Yang, Osmar Zaiane, Hi-Gcn, A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction, Comput. Biol. Med. 127 (2020).

[21] S.I. Ktena, Sarah Parisot, Enzo Ferrante, Rajchl Martin, Matthew Lee, Ben Glocker, Daniel Rueckert, Metric learning with spectral graph convolutions on brain connectivity networks, Neuroimage 169 (2018) 431–442.

[22] Xiaoxiao Li, Yuan Zhou, Nicha C. Dvornek, Muhan Zhang, Juntang Zhuang, Pamela Ventola, James S. Duncan, Pooling regularized graph neural network for fmri biomarker analysis, International Conference on Medical Image Computing and Computer-Assisted Intervention 12267 (2020) 625–635.

[23] Dongren Yao, Jing Sui, Mingliang Wang, Erkun Yang, Yeerfan Jiaerken, Na Luo, Pew-Thian Yap, Mingxia Liu, Dinggang Shen, A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity, IEEE Trans. Med. Imag., 40(2021), 1279-1289.

[24] A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F.X. Castellanos, K. Alaerts, J.S. Anderson, M. Assaf, S.Y. Bookheimer, M. Dapretto, B. Deen, S. Delmonte, I. Dinstein, B. Ertl-Wagner, D.A. Fair, L. Gallagher, D.P. Kennedy, C.L. Keown, C. Keysers, J. E. Lainhart, C. Lord, B. Luna, V. Menon, N.J. Minshew, C.S. Monk, S. Mueller, R.-A. Müller, M.B. Nebel, J.T. Nigg, K. O'Hearn, K.A. Pelphrey, S.J. Peltier, J.D. Rudie, S. Sunaert, M. Thioux, J.M. Tyszka, L.Q. Uddin, J.S. Verhoeven, N. Wenderoth, J. L. Wiggins, S.H. Mostofsky, M.P. Milham, The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism, Mol. Psychiatr. 19 (2014) 659–667.

[25] Zhen Zhou, Xiaobo Chen, Yu Zhang, Dan Hu, Lishan Qiao, Renping Yu, , Pew-Thian Yap, Gang Pan, Han Zhang, Dinggang Shen, A toolbox for brain network construction and classification (BrainNetClass), Hum. Brain Mapp. 41 (2020) 2808–2826

[26] Alexandre Abraham, Michael P. Milham, Adriana Di Martino, R. Cameron Craddock, Dimitris Samaras, Thirion Bertrand, Gael Varoquaux, Deriving reproducible biomarkers from multi-site resting-state data: an Autism-based example, Neuroimage 147 (2017) 736–745