Machine Learning Interpretability Methods to Characterize Brain Network Dynamics in Epilepsy

Dipak P. Upadhyaya, MPH[1], Katrina Prantzalos, MS[1], Suraj Thyagaraj, PhD[2], Nassim Shafiabadi, MD[1,2], Guadalupe Fernandez-BacaVaca, MD[2], Subhashini Sivagnanam, MS[3], Amitava Majumdar, PhD[3], Satya S. Sahoo, PhD[1,2]

[1] Department of Population and Quantitative Health Sciences, Case Western Reserve University School of Medicine, Cleveland, OH, USA; [2] Department of Neurology, University Hospitals Cleveland Medical Center, Cleveland, OH, USA; [3] San Diego Supercomputer Center, University of California, San Diego, CA, USA

Abstract

The rapid adoption of machine learning (ML) algorithms in medical disciplines has raised concerns about trust and the lack of understanding of their results. Efforts are being made to develop more interpretable models and establish guidelines for transparency and ethical use, ensuring the responsible integration of machine learning in healthcare. In this study we use two ML interpretability methods to gain insights into the dynamics of brain network interactions in epilepsy, a neurological disorder increasingly viewed as a network disorder that affects more than 60 million persons worldwide. Using high-resolution intracranial electroencephalogram (EEG) recordings from a cohort of 16 patients, combined with high-accuracy ML algorithms, we classify EEG recordings into binary classes of seizure and non-seizure, as well as multiple classes corresponding to different phases of a seizure. This study demonstrates, for the first time, that ML interpretability methods can provide new insights into the dynamics of aberrant brain networks in neurological disorders such as epilepsy. Moreover, we show that interpretability methods can effectively identify key brain regions and network connections involved in the disruptions of brain networks, such as those that occur during seizure events. These findings emphasize the importance of continued research into the integration of ML algorithms and interpretability methods in medical disciplines and enable the discovery of new insights into the dynamics of aberrant brain networks in epilepsy patient.

1. Introduction

Brain network dynamics play a critical role in various cognitive functions of the brain. Understanding changes in these network dynamics has been crucial in exploring disease mechanisms in neurological disorders like epilepsy [1]. Epilepsy is a serious neurological disease, affecting over 60 million individuals worldwide. Over 40% of these patients suffer from seizures that cannot be controlled with traditional anti-epileptic medications. Epilepsy has significant detrimental effects on quality of life (QoL) of patients, imposes substantial economic burden, with increased risk of mortality including poorly understood events like Sudden and Unexpected Death in Epilepsy (SUDEP). It is, therefore, vital to systematically characterize changes in brain network dynamics during epileptic seizure to achieve multiple clinical objectives, including seizure prediction [2], precise localization of seizure onset zone to enhance surgical outcomes in patients refractory to medication[3], and identifying SUDEP biomarkers [4].

There is limited understanding of the underlying principles of how changes in interactions of neurons lead to the creation of coordinated complexes that generate seizure activities; recent studies show a growing inclination towards employing rigorous mathematical models based on algebraic topology method. Algebraic topology models higher-order brain interaction between various brain regions using a generalization of the graph known as a simplex, which uses robust theoretical tools for analysis and enables the precise characterization of epilepsy networks through topological data analysis (TDA). In our previous work [5], we developed TDA methods to characterize epilepsy seizure networks using a visual representation called a persistent diagram, which systematically tracks the formation and disintegration of high dimensional topological structures. These visual representations are challenging to analyze using quantitative methods and provide limited insights into the details of aberrant brain network dynamics in epilepsy patients.

Machine Learning Framework for Brain Network Analysis

To address this challenge, we developed a machine learning (ML) framework (**Figure 1**) for the analysis of topological structures derived from high-resolution EEG data. The data is recorded using electrodes implanted in the brain using the stereotactic placement scheme to record EEG (as also called SEEG) [6]. The resulting intracranial EEG recordings are considered the "gold standard" for pre-surgical evaluation as the recordings provide high-quality signals recorded directly from brain regions, without interference from intermediate layers that often affect other types of recordings such as scalp EEG data [4]. These multi-channel signals correspond to electrical activity from different brain regions, and various methods have been developed to characterize the coupling between different signal channels during seizure-related events [7]. One widely used method is the derivation of the nonlinear regression coefficient h^2 by Pijn et al., which characterizes the coupling between two EEG signals during seizure activity. The nonlinear correlation measures are presented as a matrix (M_{ij}) showing the signal channel coupling and directionality weights [8]

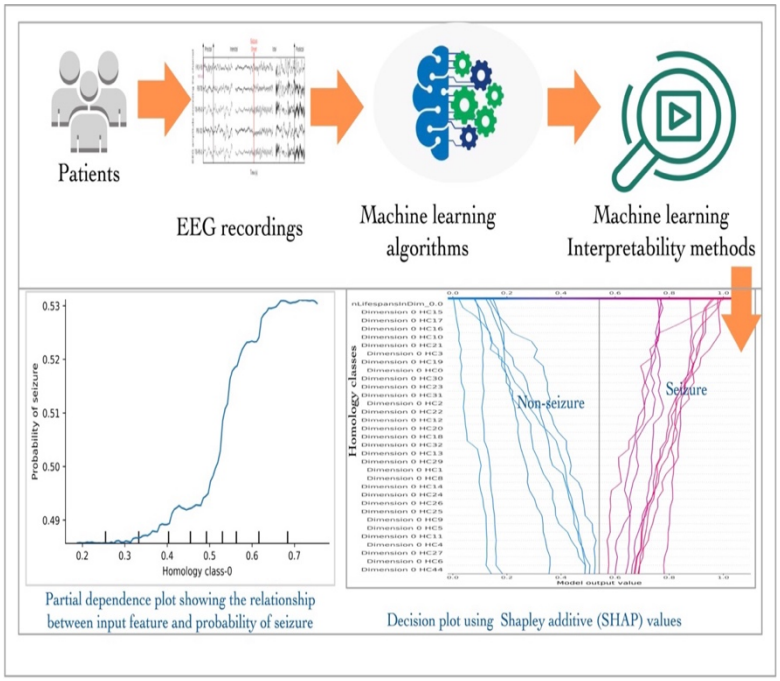

Figure 1: An overview of the machine learning interpretability framework developed within this study. In this work, we use Shapley Additive Value (SHAP) and Partial Dependence Plot (PDP) to emphasize the importance of brain interaction patterns in characterizing brain states.

In this study, we use all-to-all connected subgraphs or cliques are used as the primary construct for computing topological features. Cliques represent interacting brain structures that coordinate participation in specific brain activities, making them well-suited for studying co-activated brain structures during various epilepsy seizure events and characterizing the evolution of network structures. Therefore, topological structures generated using persistent homology (TDA method that can be used to identify cliques and other homological invariants in networks) can be used as input for ML algorithms. ML algorithms have been shown to be effective in characterizing brain states through the analysis of EEG data in various applications [9]. These algorithms excel at capturing complex patterns and relationships in EEG data that may not be apparent to human experts and have shown potential in identifying various brain states and diagnosing epilepsy [10, 11]. Recent studies have employed ML algorithms to identify brain states and interaction patterns between different brain regions. Building upon these studies, we have applied both classical and DNN algorithms to SEEG data for the following tasks:

1. **Binary classification task:** Distinguishing brain interaction patterns recorded during seizures (ictal events) and those recorded during baseline events (e.g., sleep, awake)
2. **Multiclass classification task:** Identifying brain interaction patterns recorded during different phases of epileptic seizures, including seizure onset, ictal phase 1 (first spread of seizure to additional brain regions), ictal phase 2, etc.

One key impediment to utilizing the results generated by ML algorithms in biomedical research is the limited support for interpretability and explainability. In this study, we apply two ML interpretability methods to characterize the contributions of brain patterns (represented by topological structures) in

classifying brain states using high-accuracy ML algorithms. In Section 2, we describe the implementation details of the ML algorithms for the two classification tasks listed above. In section 3, we discuss the results from this classification and subsequent interpretability analysis. Finally, in section 4 we provide discussion and conclusions from these studies.

Interpretability Methods for ML Algorithms used to Analyze EEG data. Interpretability methods are essential for understanding the variation in ML algorithm outputs with changes in input features, while explainable methods are employed to characterize the dynamics of ML algorithms [12]. In this paper, we utilized two post-hoc model-agnostic interpretability methods to analyze the contribution of different categories of brain interaction patterns represented by algebraic topology structures, that were used as input features for the three ML models. The first interpretability method used in this study is the Shapley Additive Value (SHAP), and the second method is the Partial Dependence Plot (PDP).

The SHAP method is based on cooperative game theory and provides an explanation of the model's complexity [13]. It provides interpretability to the model's predictions by employing a linear approximation based on the original machine learning methods. This allows for a better understanding of how individual features contribute to the model's output. SHAP computes the average marginal contribution of features across all possible combinations, shedding light on their individual impact [13]. The second interpretability method employed in this study is PDP, a powerful model-agnostic graphical method commonly used to describe the relationship between input features and model outcomes at a global level. PDP is employed as a statistical estimator of the ground truth to interpret the data generated by the model. It is particularly useful in detecting non-linear relationships between features and the outcome that may not be evident from simple correlation analysis or other exploratory data analysis techniques. It serves as a powerful tool for model interpretation, providing valuable insights into how the model generates results.

1.1 Background

Topological data analysis for characterizing epileptic seizure networks. Current network analysis methods using graph models, where vertices represent brain structures and edges represent brain interactions, are limited to modeling binary interactions between two vertices. However, in neurological disorders like epilepsy, there are often simultaneous interactions between three or more brain structures. The use of TDA methods, such as persistent homology, to characterize topological structures in noisy datasets has gained significant attention in brain network studies. In TDA, data can be encoded as a simplicial complex, a generalization of a graph comprised of simplices that describes multidimensional interactions. A k-simplex is the convex hull of k+1 interacting nodes, and the value of k is known as the dimension of the simplex. TDA analyzes simplicial complexes to identify topological structures, called homology classes, which are the boundary of a collection of k-simplices that form a closed structure (such as a loop or a tunnel). The dimension of a homology class is equivalent to the dimension of the simplices forming the closed structure. A dimension 0 homology class represents the vertices in a network; a dimension 1 homology class represents the boundary of a 1-dimensional closed structure; a dimension 2 homology class represents the boundary of a 2-dimensional closed structure with additional dimensions corresponding to higher dimensional structures.

Persistent homology tracks the formation and termination of these homology classes across a filtration, such as thresholds for correlations in brain activity. In terms of brain networks, the lifespan (filtration of termination minus the filtration of formation) of these homology classes can be used to describe the strength of interaction between brain regions. Homology classes with larger lifespans are assumed to describe more fundamental features of the complex; however, short-lived homology classes may be equally as important in characterizing networks [14]. In this work we use persistent homology to identify the lifespans of homology classes that represent important topological structures in brain networks. We use these lifespans as input into machine learning models and apply interpretability methods to these models to determine the use of homology classes and their lifespans in characterizing dynamic brain networks.

Clinical goals in studying epileptic seizure networks. A systematic study of brain connectivity disorders and dysfunctional brain system dynamics can provide crucial insights into higher-order brain functions, including consciousness, attention, and working memory. Advances in large-scale network analysis using graph models and the identification of intrinsic connectivity networks (ICN) have significantly contributed to the study of dysfunctional brain systems. The changes in these ICNs during clinically defined dysfunctional brain states can be correlated with aberrant alterations in brain functions. Recent studies have shown the deactivation of brain regions within the default mode network (DMN) during epileptic seizures, accompanied by a loss of consciousness [15]. Therefore, a systematic study of aberrant brain network dynamics holds significant clinical relevance for physicians, epilepsy patients, and their families.

Related Work. As part of our preliminary work, we analyzed brain interaction data recorded from intracranial electrodes in a cohort of five patients with refractory epilepsy [16]. We computed directed graphs using the nonlinear correlation coefficient h^2, with vertices representing intracranial electrode contacts and edges representing correlation. Persistent homology was applied to these graphs to derive simplicial complexes during seizure events. Our findings revealed that topological structures formed during seizure onset and ictal 1 phases exhibited longer persistence compared to structures formed during subsequent events (e.g., ictal 2, ictal 3 events etc.) [16]. These results support our hypothesis that topological structures serve as robust discriminating features for characterizing pathological brain states. Similarly, other studies have also employed TDA in conjunction with ML algorithms for binary classification tasks, distinguishing between seizure and non-seizure event. In our recent work [17], we utilized TDA for multiclass classification of EEG data across different stages of epileptic seizures, as discussed in the previous section.

2. Methods

2.1 Dataset and Patient Cohort

We analyzed data from a cohort of 16 patients with refractory epilepsy, who underwent presurgical evaluation at the Epilepsy Center of University Hospitals Cleveland Medical Center, a Level 4 epilepsy center [17]. The cohort was selected to determine the suitability of surgical intervention for disrupting the propagation networks of epileptic seizures. SEEG recordings corresponding to two randomly selected seizure events were selected for each patient. Recordings from one non-seizure period of asleep and one non-seizure period of awake were also selected for each patient, matched in duration to the longer of the two seizure events. Bidirectional network graphs for one-second epochs of SEEG recordings were computed using the nonlinear correlation coefficient h^2 Pijn [8]. All correlations were processed using the Neuro-Integrative Connectivity (NIC) tool [5]. Persistent homology was applied to each graph to determine the lifespan of homology classes in each one-second epoch. The lifespan values of all homology classes related to one epoch were stored in a list. Each epoch's list was padded with 0s, representing homology classes that had a lifespan of 0, to ensure that all lists were the same length. Each list was labeled according to its associated period (seizure or non-seizure in binary classification or period of the seizure in multiclass classification). These lists were used as input into various machine learning models.

2.2 Data Augmentation for Imbalanced Data

The original dataset encompassed data from six distinct classes with the minority class (onset, with 294 epochs) comprising 9% and the majority class (ictal 3, with 755 epochs) comprising 23% of the total data. This led to an imbalanced dataset. To mitigate this imbalance issue and rectify the uneven distribution of seizure states within each class, we employed Synthetic Minority Oversampling Techniques (SMOTE), a data augmentation approach to address class label imbalanced in the study datasets [18]. Imbalanced data, where the number of samples from each class differs significantly, is a common challenge in real-world scenarios such as healthcare. Imbalanced classes compromise classifier performance and can result in biased predictions and increased error costs, particularly in healthcare settings. To address the effects of imbalanced data on classifier performance, we employed SMOTE, a data level approach to data augmentation in which the minority classes were oversampled to balance the data. SMOTE uses

randomized interpolation of neighboring minority labels from a set of instances representing the minority classes to generate new instances of minority classes, enhancing classifier generalizability. It does not introduce new information or variation but rather generates new data points using the K-nearest method from the feature space.

2.3 Machine Learning Models

In the study, three machine learning algorithms were used to characterize brain states and understand the interactions between different brain regions: Support Vector Machine (SVM), Random Forest (RF), and Deep Neural Network (DNN).

Support Vector Machine (SVM) is a supervised learning algorithm that aims to find the optimal classification hyperplane to distinguish the different categories of data. SVM can handle both linearly separable and non-linearly separable data by transforming the input data into a higher-dimensional space using a kernel function. It finds the optimal hyperplane by maximizing the margin between classes. SVM has advantages such as being less prone to overfitting, working well with high-dimensional data, and handling binary and multi-class classification problems.

Random Forest (RF) is an ensemble learning algorithm that combines multiple decision trees to make predictions. RF introduces randomness by training each tree on a random subset of the data and features, reducing overfitting. The algorithm creates a set of decision trees by recursively partitioning the data based on feature values. RF improves accuracy by taking a majority vote on predictions made by each tree. It can handle both categorical and continuous data, handle missing values, and identify important features.

Deep Neural Network (DNN) models are inspired by the human brain's structure and function. A DNN consists of interconnected layers of neurons that learn complex relationships between input data and output predictions. DNNs have multiple hidden layers that perform non-linear transformations on input data to extract higher-level features and patterns. Each hidden layer's output is used as input for the next layer, allowing the network to learn increasingly complex representations. DNNs have achieved state-of-the-art performance in various applications.

For Support Vector Machine (SVM) analysis, we utilized the SVC classifier from the scikit-learn package, employing specific hyperparameters such as regularization (C = 10), kernel type (poly), and the degree of the kernel function. The SVM model was also configured to provide probability estimates. The Deep Neural Network (DNN) architecture was implemented using the TensorFlow Keras API. It consisted of five hidden layers with a varying number of nodes, and the rectified linear unit (ReLU) activation function was applied to each layer. The model utilized the "rmsprop" optimizer, "Binary_crossentropy" loss function, and a learning rate of 0.001. In binary classification, the final output layer employed a sigmoid activation function, while in multiclass classification, the softmax activation function was used. For ensemble learning in classification tasks, Random Forest was employed. We utilized the RandomForestClassifier from the scikit-learn package, specifying the number of trees in the forest (n_estimators = 100), maximum depth (max_depth = 5), minimum number of samples required to be in a leaf node (min_samples_leaf = 1), and the minimum number of samples required to split an internal node (min_samples_split = 15). To optimize the hyperparameters for these machine learning algorithms, we employed the Keras grid search library. The objective of this study was to comprehend the underlying patterns in brain state data and validate the results across different models by comparing their performance.

2.4 Interpretability methods

In the second phase of the study, the interpretability methods were applied to the machine learning algorithm with the highest accuracy. The objective was to systematically analyze and characterize the contribution of individual input features to the classification results generated by the two ML algorithms with the highest accuracy. We used the SHAP decision plot and SHAP summary plot (also known as a waterfall plot) as interpretability methods to visualize the results of the machine learning models. These

methods provide insights into the model's decision-making process and the impact of different features on the predictions.

1. SHAP Decision Plot: The SHAP decision plot helps in understanding how complex models make predictions and demonstrates the influence of various features on individual predictions. It allows for the examination of the effects of multiple features on the model's output for a specific instance. By analyzing the SHAP decision plot, one can gain a deeper understanding of how the model arrives at its predictions and the relative importance of different features for each prediction.
2. SHAP Summary Plot (Waterfall Plot): The SHAP summary plot, or waterfall plot, provides an overview of feature importance and shows how the combination of features affects the model's predictions. It illustrates the contribution of each feature by adding or subtracting from a baseline value (e.g., average prediction). The waterfall-like visualization helps in comprehending the overall impact of different features on the model's output and their interactions.

We also used PDP as interpretability method, which holds all other features constant and varies the feature of interest. The resulting plot illustrates the effect of changing the value of the feature of interest on the predicted outcome while averaging the effects of all other features. Typically, the plot displays the predicted outcome on the y-axis and the values of the feature of interest on the x-axis. PDP helps to uncover insights about feature importance, linearity, interactions, and potential biases, enhancing transparency and understanding of model behavior.

3. Results

In this section, we present the results obtained from applying interpretability methods to our machine learning algorithms for the binary classification task of distinguishing between seizure brain states and non-seizure brain states, as well as the multiclass classification task of distinguishing between different phases of seizures. As described in Section 1, we utilized a standardized approach to evaluate the classification performance. We found that these models successfully distinguished seizure states (Table 1). In the following subsections, we provide a detailed description of the results obtained from these interpretability methods for each classification task: binary classification (Section 3.1.1) and multiclass classification (Section 3.1.2).

Table 1: Accuracy values of binary and multiclass classification across all three algorithms for the complete study cohort

Machine learning algorithms	**Binary Classification** (Seizure and non-seizure states)	**Multiclass classification** (States of seizure progressions)
Support Vector Machine	0.81	0.64
Deep Neural Network	0.82	0.64
Random Forest	0.88	0.74

3.1 Interpretability Methods for ML algorithms: In this study, we employed two interpretability methods, specifically SHAP and PDP, on the Random Forest algorithm, which yielded the highest accuracy values of 0.88 and 0.74 for binary and multiclass classification, respectively. We determined the feature importance scores for models applied to the whole cohort as well as models for each individual patient in our cohort.

3.1.1. Interpretability results for binary classification task. We first applied interpretability methods to ascertain the contribution of homology classes and their lifespans to the classification of seizure and non-seizure states in a Random Forest model. We show that the lifespan of homology classes in dimension 0, representing the filtration (strength of correlation) at which an electrode becomes involved in any sort of connection (pairwise or multidimensional), are consistently the most distinguishing features in seizure or non-seizure classification in single patient analyses or across all patients (**Figure 2**). We also show that there is some importance of the dimension 1 homology class lifespans. A homology class in dimension 1 is the boundary of a collection of 1-simplices (pairwise connections) that form a closed path (such as a loop or a cycle). This means that, for each dimension 1 homology class, there is a set of electrodes for which

there is a continuous and closed path of pairwise interaction; however, at least one pair of electrodes in the interacting set does not directly interact. The lifespan of a dimension 1 homology class thus represents how much more interaction is needed for an existing path of pairwise interactions to interact with other electrodes. Our results indicate that the lifespans of these lower-dimensional interactions are the most influential in classifying seizure and non-seizure states.

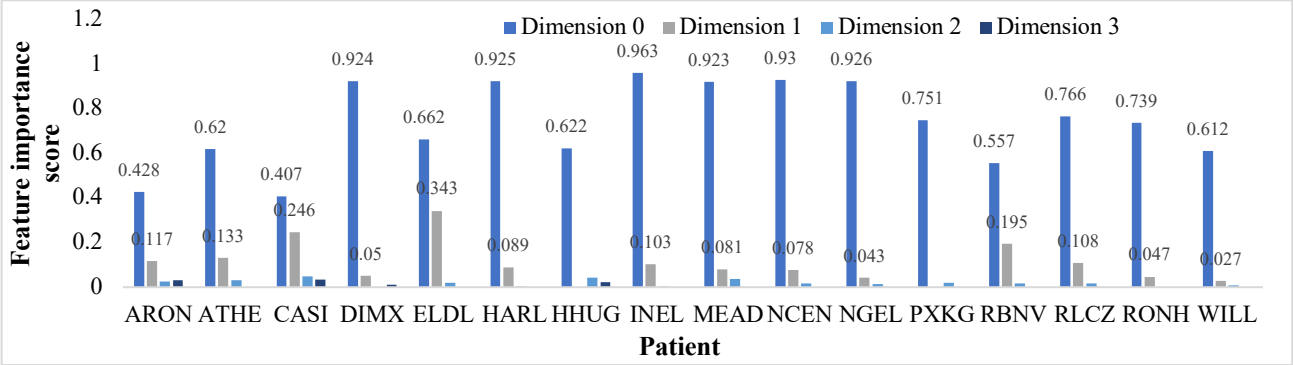

Figure 2: Feature importance scores for homology classes along each dimension

To identify the impact of different homology classes on the classification of seizure and non-seizure events in EEG samples, we selected the first one-second epoch from seizure and non-seizure periods from a randomly selected patient and visualized them using a SHAP decision plot (**Figure 3**). In this plot, homology classes are labeled with their dimension and their order in the lifespan list (homology class 1 (HC1), HC2, etc.), and are organized in decending order of there importance as determined by their SHAP values. We note that the identifiers associated with each homology class is generated by the GUDHI TDA tool [19] integrated in to the NIC tool.

SHAP values indicate the disparity between the predicted and actual outputs of the machine learning algorithms [13]. For both seizure and non-seizure states, the decision plot starts at the base of the y-axis, representing the expected value generated by the machine learning algorithm. Step by step, the SHAP values corresponding to each feature are added to the expected value, resulting in the final classification outcome predicted by the algorithm.

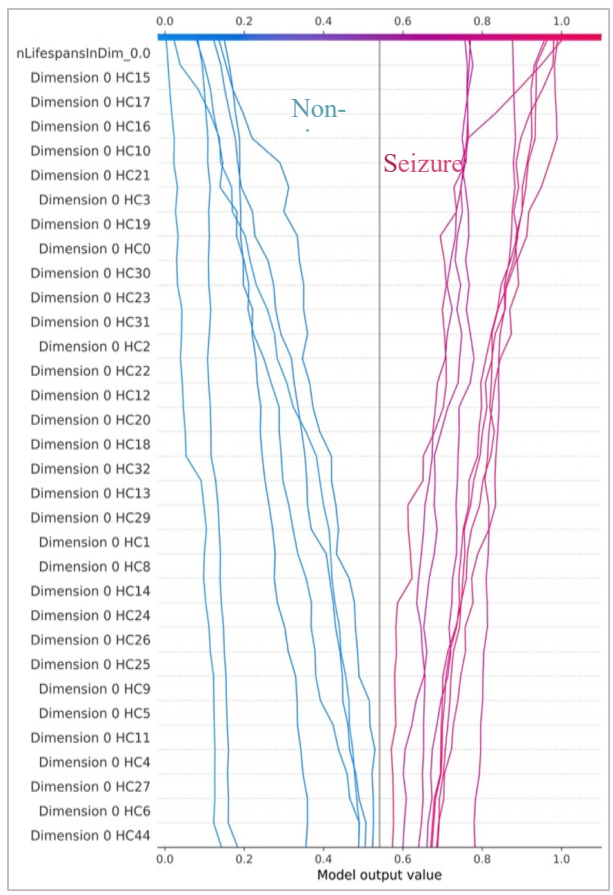

Figure 3: A decision plot visualization of SHAP feature importance score computed for different homology classes across different dimensions.

The findings illustrated in **Figure 3** highlight that a significant number of homology classes in dimension 0 play a crucial role in accurately classifying seizure and non-seizure events. This further emphasizes the importance of lower dimensional homology classes and their lifespans in characterizing seizure and non-seizure states. Through **Figure 3**, we are able to identify several dimension 0 homology classes which are most influential in classifying seizure and non-seizure states, such as

homology class 15, 16, and 17 (identifiers generated by the GUDDI tool). The importance of the difference in lifespan values for these homology classes can be seen further when looking at persistence diagrams of each of the corresponding one-second epochs (**Figure 4**). For example, using the same epochs as selected for SHAP decision plot visualization, we suggest that asleep non-seizure periods might be classified by recognizing higher lifespans for homology classes 15, 16, and 17 compared to seizure onset; however, awake non-seizure periods might be characterized by much lower lifespans for homology classes 15, 16, and 17 (identifiers generated by the GUDHI tool).

In addition to SHAP, we also applied PDP to our random forest model. PDP illustrates the functional relationship between homology classes, dimensions, and seizure probability. While SHAP identified dimension 0 homology classes to be the most essential for classification, PDP reveals some relationships in higher dimensions to have significant impact on seizure classification (**Figure 5**). Our results show that dimension 0 homology class 0, dimension 1 homology class 1, and dimension 5 homology class 0 have a linear effect on seizure probability, with increasing dimensions and classes lifespan values corresponding to higher probabilities. However, dimension 2 homology class 2, dimension 3 homology class 1, and the maximum dimension exhibit nonlinear effects. The relationship between these features and seizure probability does not follow a linear trend, but rather shows fluctuations at specific threshold values. In **Figure 5**, we observe an initial increase in seizure probability as the lifespan value of dimension 2 homology class 0 reaches 0.03, followed by a substantial decrease beyond this threshold. These findings emphasize the importance of considering both dimension and homology class in analyzing EEG recordings and highlight the complex nature of their relationship.

3.1.2. Interpretability results for multiclass classification task.

In order to interpret the random forest algorithm results for our multiclass classification of seizure periods, we created a waterfall plot (**Figure 6**) to demonstrate the SHAP feature importance scores

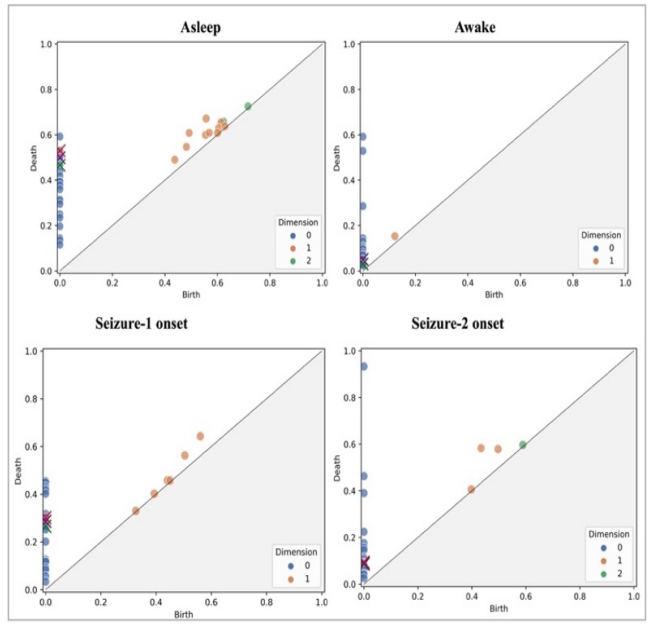

Figure 4: Scatter plots of three most distinguishing homology classes for a patient (persistence diagrams) corresponding to asleep, awake, seizure one onset, and seizure two onset states. Homology classes marked by colors red, green, and purple in dimension 0 were the most influential features.

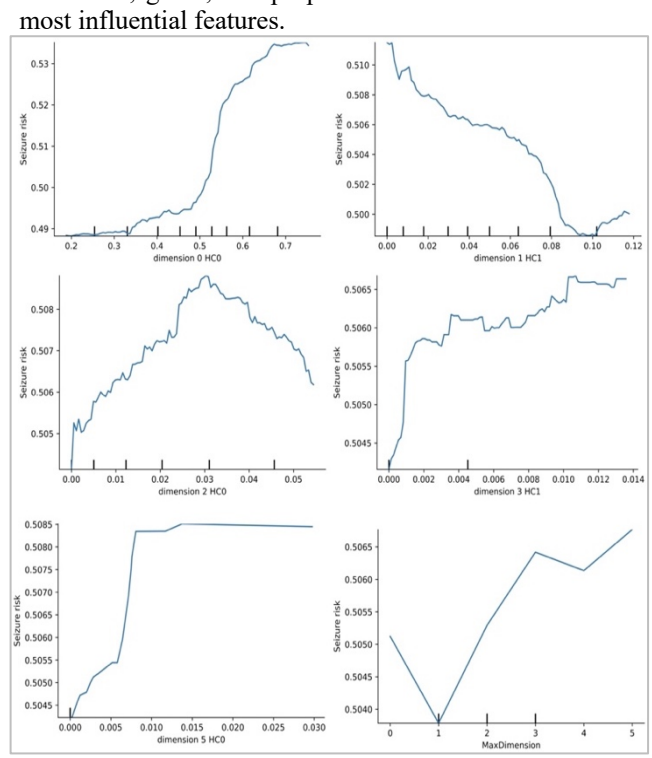

Figure 5: The partial dependence plot (PDP) for five homology classes with highest feature importance scores together with partial dependency score of the number of dimensions.

of homology classes in each dimension. Similar to binary classification results, our multiclass classification results reveal that lifespans of homology classes in dimension 0 (representing the lowest filtration at which an electrode interacts with any other electrodes) make significant contributions to classification of seizure periods. Additionally, we find that dimension 1 homology class lifespans are of high importance when distinguishing the third ictal period, suggesting that dimension 1 homology classes carry valuable information for distinguishing the ictal 3 state from other seizure states.

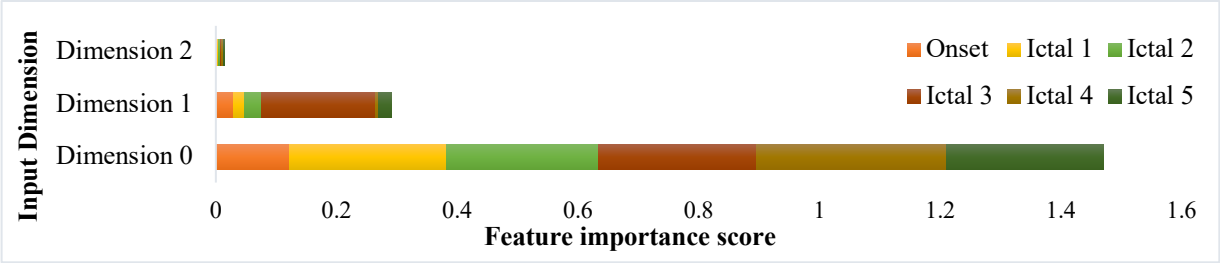

Figure 6: Waterfall plot of the SHAP feature importance scores for multiclass classification task.

4. Discussion and Conclusions

The results of the study demonstrate that ML interpretability methods can effectively identify brain network connections involved in seizure events. This study provides new avenues for understanding complex brain network dynamics in epilepsy and other disorders, highlighting the potential of ML interpretability methods to provide insights into neurological disorders and brain network dynamics. Our findings emphasize the importance of continued research into the integration of ML algorithms and interpretability methods in medical disciplines, enabling more accurate and personalized diagnosis and treatment options for patients. This study also contributes to the growing body of knowledge regarding the use of ML in medical applications, providing new avenues for understanding complex brain network dynamics in epilepsy and other disorders.

Acknowledgements

This research was funded in part by the grants from the US National Institutes of Health (NIH): U24EB029005, R01DA053028, the US Department of Defense (DoD) grant#W81XWH2110859, Dravet Syndrome Foundation (DSF), and the Clinical and Translational Science Collaborative of Cleveland, which is funded by the NIHNCATS, Clinical and Translational Science Award grant, UL1TR002548. The content is solely the responsibility of the authors and does not necessarily represent the official views of the respective funding agencies.

References

1. Saggar, M., et al., *Towards a new approach to reveal dynamical organization of the brain using topological data analysis.* Nat Commun, 2018. **9**(1): p. 1399.
2. Kuhlmann, L., et al., *Seizure prediction - ready for a new era.* Nat Rev Neurol, 2018. **14**(10): p. 618-630.
3. Burns, S.P., et al., *Network dynamics of the brain and influence of the epileptic seizure onset zone.* Proc Natl Acad Sci U S A, 2014. **111**(49): p. E5321-30.
4. Devinsky, O., et al., *Sudden unexpected death in epilepsy: epidemiology, mechanisms, and prevention.* Lancet Neurol, 2016. **15**(10): p. 1075-88.
5. Sahoo, S.S., et al., *NeuroIntegrative Connectivity (NIC) Informatics Tool for Brain Functional Connectivity Network Analysis in Cohort Studies.* AMIA Annu Symp Proc, 2020. **2020**: p. 1090-1099.
6. Talairach, J., Tournoux, P., *Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System - an Approach to Cerebral Imaging.* Thieme Medical, 1988.
7. Bullmore, E. and O. Sporns, *Complex brain networks: graph theoretical analysis of structural and functional systems (vol 10, pg 186, 2009).* Nature Reviews Neuroscience, 2009. **10**(4).

8. Greicius, M.D., et al., *Functional connectivity in the resting brain: A network analysis of the default mode hypothesis.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(1): p. 253-258.
9. Aggarwal, S. and N. Chugh, *Review of Machine Learning Techniques for EEG Based Brain Computer Interface.* Archives of Computational Methods in Engineering, 2022. **29**(5): p. 3001-3020.
10. Rasheed, K., et al., *Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A Review.* IEEE Rev Biomed Eng, 2021. **14**: p. 139-155.
11. Roy, Y., et al., *Deep learning-based electroencephalography analysis: a systematic review.* J Neural Eng, 2019. **16**(5): p. 051001.
12. Russell, S., Norvig, P., *Artificial intelligence: a modern approach*. 4th edition ed. 2020: Pearson.
13. Lundberg, S.M. and S.I. Lee, *A Unified Approach to Interpreting Model Predictions.* Advances in Neural Information Processing Systems 30 (Nips 2017), 2017. **30**.
14. Sizemore, A.E., et al., *The importance of the whole: Topological data analysis for the network neuroscientist.* Netw Neurosci, 2019. **3**(3): p. 656-673.
15. Danielson, N.B., J.N. Guo, and H. Blumenfeld, *The default mode network and altered consciousness in epilepsy.* Behavioural Neurology, 2011. **24**(1): p. 55-65.
16. Zhang, J., et al., *Characterizing Brain Network Dynamics using Persistent Homology in Patients with Refractory Epilepsy.* AMIA Annu Symp Proc, 2021. **2021**: p. 1244-1253.
17. Prantzalos, K., Bauman, R., Shafiabadi, N., Thyagaraj, S., Gurski, N., Miller, J., Fernandez-BacaVaca, G., Sahoo, S.S., *Distinguishing Aberrant Brain Network States using Persistent Homology in a Machine Learning Workflow*. 2023, Department of Population and Quantitative Health Sciences, Case Western Reserve University School of Medicine, Cleveland, OH, USA.
18. Chawla, N.V., et al., *SMOTE: Synthetic minority over-sampling technique.* Journal of Artificial Intelligence Research, 2002. **16**: p. 321-357.
19. Maria, C., Boissonnat, J.-D., Glisse, M. & Yvinec, M. *GUDHI library. https://project.inria.fr/gudhi/software/ (2014)*