Explainable Disentangled Representation Learning of Recurring Brain Activation Patterns via Variational Autoencoders

Saheed Faremi[1,*]

[1]School of Computer Science & Information Technology, University College Cork, Ireland

Abstract

Electroencephalography (EEG) microstates are quasi-stable topographic patterns of brain electrical activity lasting 60–120ms, believed to reflect fundamental modes of neural processing. Current methods for identifying these patterns rely on rigid clustering approaches that conflate distinct neurophysiological factors, lack temporal modelling, and offer limited interpretability. This limits their utility for understanding brain disorders, where predicting transitions between brain states could have significant clinical value. This research proposes a framework combining convolutional variational autoencoders with Gaussian mixture model priors and multi-objective optimisation to learn disentangled representations of recurring brain activation patterns. By separating latent factors while preserving temporal transition dynamics, the framework aims to make individual brain states both interpretable and generalisable across subjects. Explainability is embedded through latent traversals, feature attribution, and counterfactual explanations to ensure the learned representations are transparent to neuroscientists and clinicians. Early investigations on the LEMON dataset show promise in recovering structured latent representations consistent with known microstate topographies. These preliminary findings motivate the proposed multi-objective framework as a pathway toward representational learning of recurring brain activation patterns, such that subject-specific patterns are generalisable to group patterns through an interpretable and explainable latent space. This work introduces an architecture for identifying recurring brain activation patterns and predicting subsequent brain states, with applications to brain disorder diagnosis.

Keywords

Explainable Artificial Intelligence, EEG Microstates, Variational Autoencoders, Disentangled Representations, Multi-Objective Optimisation,

1. Introduction

Electroencephalography (EEG) captures brain electrical activity at millisecond resolution, revealing that scalp topographies remain quasi-stable for brief periods (60–120 ms) before transitioning abruptly to a different configuration. These quasi-stable states, termed *EEG microstates*, have been described as the "atoms of thought" [1] and are widely studied as potential biomarkers for cognitive states and neuropsychiatric conditions [2, 3]. Canonical microstate analysis identifies four to seven recurring topographic maps (typically labelled A, B, C, D) using modified k-means clustering at Global Field Power (GFP) peaks [4]. Despite their utility, conventional methods suffer from critical limitations. First, modified k-means and its variants assign each time point to a single cluster, conflating multiple neurophysiological factors such as anterior-posterior gradients, laterality, and frequency-band dominance into a single label. This precludes understanding *why* a given topography belongs to a particular microstate class. Second, clustering operates independently at each GFP peak, discarding the temporal transition dynamics that carry rich information about brain state sequences [5]. Third, the resulting explanations are opaque: a cluster centroid provides limited insight into the neurophysiological processes generating that pattern. These limitations motivate a fundamentally different approach rooted in deep generative modelling and Explainable AI (XAI).

Late-breaking work, Demos and Doctoral Consortium, colocated with the 4th World Conference on eXplainable Artificial Intelligence: July 01–03, 2026, Fortaleza, Brazil

*Corresponding author.

126100912@umail.ucc.ie (S. Faremi)

0009-0007-8428-7019 (S. Faremi)

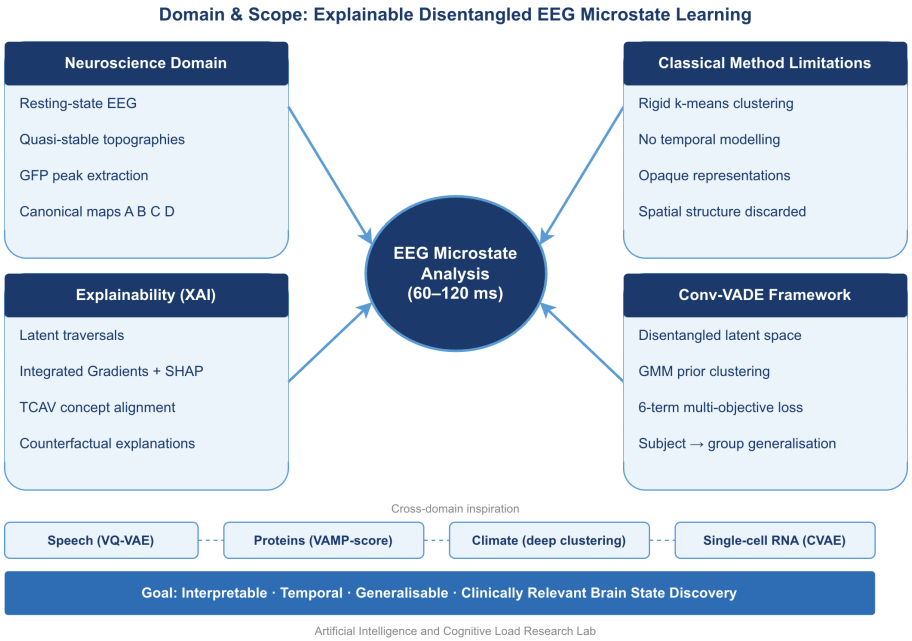

Figure 1: Domain and Scope

Variational Autoencoders (VAEs) [6] learn continuous latent representations that can capture multiple independent factors of variation simultaneously. When combined with structured priors such as Gaussian Mixture Models (GMMs), VAEs can perform clustering and representation learning jointly [7]. Crucially, disentanglement techniques originating from the computer vision literature [8, 9] enable individual latent dimensions to correspond to interpretable factors, a property directly aligned with XAI objectives. This research proposes a Convolutional Variational Deep Embedding (Conv-VADE) multi-objective framework for disentangled representational learning of recurring EEG microstate patterns from spatial topographic maps. By learning representations where individual latent dimensions correspond to identifiable neurophysiological factors, the approach aims to transform microstate analysis from black-box clustering into transparent, explainable brain state discovery. This is based on the research question: Can variational deep embedding methods learn recurring patterns from EEG spatial topographic maps rather than single vector representations, such that subject-specific representations are generalisable to group-level signatures, enabling temporal sequencing of brain states that supports the diagnosis of brain disorders, through a framework that is explainable and interpretable? The remainder of this proposal is organised as follows: Section 2 reviews related work in EEG microstate analysis and deep generative modelling; Section 3 presents the research questions, hypothesis, and objectives; Section 4 describes the proposed methodology; Section 5 reports preliminary results; and Section 6 outlines expected next steps and final contributions.

2. Related Work

This section reviews five complementary research threads relevant to this doctoral work: classical EEG microstate analysis, variational autoencoders and disentanglement, variational deep embedding and clustering, explainability for generative models, and cross-domain recurring state discovery. Classical microstate analysis employs modified k-means [4] or atomise-and-agglomerate hierarchical clustering (AAHC) [10] to identify recurring topographic patterns at GFP peaks. While effective for canonical map recovery, these methods treat topographies as monolithic patterns without decomposing them into constituent neurophysiological factors. Recent extensions incorporate Hidden Markov Models (HMMs) to capture temporal dynamics of brain states [11], but the state representations themselves remain unstructured and difficult to interpret. Polarity invariance the biophysical property that topographies

of opposite polarity reflect the same underlying source configuration is typically handled through absolute value transformations that discard potentially informative asymmetry.

VAEs [6] provide a principled framework for learning latent representations of complex data. The β-VAE [8] introduced controllable disentanglement by up-weighting the KL divergence term, while β-TC-VAE [9] refined this by penalising total correlation specifically. Quantitative disentanglement evaluation metrics including DCI [12], MIG [9], and SAP [13] exist but require known ground-truth factors a challenge for EEG where such factors must first be defined.VaDE (Variational Deep Embedding) [7] integrates GMM priors into the VAE latent space, enabling joint clustering and representation learning. This architecture is particularly relevant for microstates, where both cluster identity and within-cluster variation carry scientific meaning. However, VaDE and its variants have not been applied to EEG data or adapted for the temporal and invariance constraints specific to brain signals. Explainability methods for time series classifiers have advanced considerably [14], with approaches including LIME adaptations [15], rigorous evaluation frameworks for time-series XAI [16], and attention-based visualisations. However, XAI for *generative* models of neural data explaining *what* the model has learned, not just *why* it classified requires different tools: latent traversals [8], concept-based explanations [17], and counterfactual generation. These methods are well-established in computer vision but have not been systematically applied to EEG representations.

The abstract problem of discovering recurring quasi-stable states in high-dimensional temporal data appears across diverse domains. Unsupervised speech recognition [18] learns discrete acoustic units from continuous signals via quantised self-supervised representation learning; Vector Quantised VAE (VQ-VAE) architectures [19] similarly learn discrete codes for recurring units. Protein conformational state analysis employs time-lagged autoencoders with Variational Approach for Markov Processes score (VAMP-score) objectives [20] to preserve slow dynamics. Climate regime detection uses deep clustering on spatial fields [21]. Single-cell RNA sequencing analysis addresses cross-subject variability through conditional VAEs [22]. These cross-domain solutions offer transferable architectural and objective-function innovations for EEG microstate analysis. Across the reviewed literature, a consistent set of limitations emerges. Classical microstate methods reduce rich spatial topographic maps to single vector representations, discarding the within-map structure that carries neurophysiological meaning. Temporal modelling approaches such as HMMs partially address transition dynamics but operate on unstructured state representations that remain opaque and uninterpretable. Cross-domain solutions demonstrate that recurring quasi-stable patterns can be learned from high-dimensional temporal data, yet none have been applied to EEG topographic maps with explicit disentanglement objectives. Critically, no existing method simultaneously addresses spatial richness, temporal coherence, explainability, and subject-to-group generalisation within a single framework. This gap motivates the proposed Conv-VADE multi-objective framework, which learns disentangled representations directly from EEG spatial topographic maps, preserves temporal transition dynamics, and produces interpretable and explainable brain state signatures generalisable across subjects. The research questions, hypothesis, and objectives that follow are directly grounded in these identified gaps and position the proposed Conv-VADE framework as a principled response to the limitations of existing approaches.

3. Research Questions, Hypothesis, and Objectives

The overarching research question guiding this doctoral work is: Can variational deep embedding methods learn recurring patterns from EEG spatial topographic maps rather than single vector representations, such that subject-specific representations are generalisable to group-level signatures, enabling temporal sequencing of brain states that supports the diagnosis of brain disorders, through a framework that is explainable and interpretable?
The following hypothesis is proposed in response to this question: If a Convolutional Variational Deep Embedding (Conv-VADE) model is trained on EEG data using a multi-objective loss function, then it will

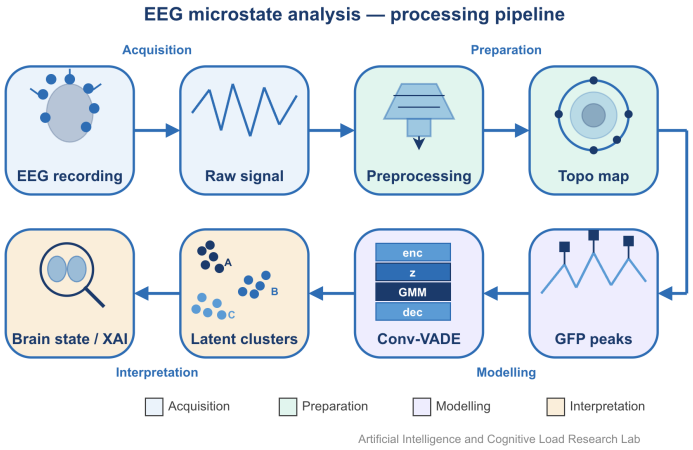

Figure 2: Overview of the Conv-VADE research framework structured around five CRISP-DM phases. The figure illustrates the flow from EEG data acquisition through preprocessing, model training, XAI integration, and evaluation, with arrows indicating iterative feedback between phases.

learn recurring patterns from spatial topographic maps such that subject-specific latent representations are generalisable to group-level signatures that are both interpretable and explainable, with initial validation on the LEMON dataset. This question and hypothesis are decomposed into the following specific objectives:

O1: To develop a Conv-VADE architecture that jointly learns disentangled latent representations and cluster assignments for EEG topographies, recovering canonical microstate maps while revealing finer-grained neurophysiological structure.

O2: To design a multi-objective optimisation framework that balances reconstruction fidelity, latent disentanglement (via total correlation penalisation), GMM clustering quality, temporal transition coherence, and polarity invariance.

O3: To develop an XAI framework comprising latent traversals, feature attribution (Integrated Gradients [23], SHAP [24]), concept alignment (Testing with Concept Activation Vectors, TCAV [17]), and counterfactual explanations specifically tailored for interpreting learned EEG representations.

O4: To establish evaluation protocols and metrics spanning reconstruction quality, disentanglement (DCI, MIG, SAP), clustering validity, temporal fidelity (transition matrix correlation, duration distribution matching), and neuroscientific validity (canonical map recovery, source localisation consistency) for comprehensive assessment of the framework.

O5: To validate the clinical relevance of disentangled microstate representations through cross-dataset generalisation and association with cognitive or clinical phenotypes.

4. Research Plan and Methodology

This research follows the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, adapted for a neuroscientific deep learning context. The five phases applied are: Business Understanding, Data Understanding, Data Preparation, Modelling, and Evaluation. An overview of the research plan is provided in Figure 2.

The goal of this research is to develop a model capable of identifying recurring patterns in EEG spatial topographic maps, such that subject-specific brain state representations are generalisable to group-level signatures, enabling temporal sequencing of brain states that supports the diagnosis of brain disorders. Conventional microstate analysis reduces rich topographic maps to single vector representations, discarding spatial structure and temporal dynamics. This research addresses that gap through a Convolutional Variational Deep Embedding (Conv-VADE) multi-objective framework

grounded in Explainable AI (XAI) principles. The following assumptions underpin this research: (1) GFP peaks are appropriate extraction points for microstate topographies; (2) the LEMON dataset is representative of healthy resting-state EEG; (3) a GMM prior is an appropriate probabilistic structure for microstate clustering; and (4) four to seven microstates is the appropriate number of classes for canonical map recovery.

4.1. Data Understanding and Preparation

EEG data is sourced from the LEMON dataset [25], a large-scale, publicly available resting-state EEG resource comprising recordings from healthy adult participants. The dataset provides high-density EEG recordings suitable for topographic microstate analysis. Each recording captures spontaneous brain electrical activity at millisecond resolution, producing spatial topographic maps across the scalp electrode array. Data is partitioned at the subject level into training, validation, and test sets, ensuring no subject appears across multiple splits. This prevents data leakage and supports the generalisation claim from subject-specific to group-level representations. EEG data undergoes standard preprocessing using MNE-Python [26]: bandpass filtering to isolate relevant frequency bands, artifact rejection via Independent Component Analysis (ICA), and re-referencing to average reference. Topographies are extracted at GFP peaks following established protocols [2]. Both the original topography $\mathbf{x}$ and its polarity-inverted version $-\mathbf{x}$ are included during training to address polarity invariance through data augmentation.

4.2. Modelling

The modelling phase is structured into three components implemented in PyTorch, with supporting libraries including NumPy, SciPy, Scikit-learn, and TensorFlow/Keras for comparative baselines. The Conv-VADE architecture employs convolutional layers operating on electrode-space representations, encoding each topography into a latent vector $\mathbf{z} \in \mathbb{R}^d$. The latent space is structured by a GMM prior:

$$p(\mathbf{z}) = \sum_{k=1}^{K} \pi_k \, \mathcal{N}(\mathbf{z} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \tag{1}$$

where K corresponds to the number of microstate classes, π_k are mixture weights, and $\boldsymbol{\mu}_k$, $\boldsymbol{\Sigma}_k$ are learnable component parameters. This formulation enables simultaneous representation learning and clustering within a single probabilistic framework. Polarity invariance is further enforced through a consistency loss that encourages $\mathbf{z}(\mathbf{x}) \approx \mathbf{z}(-\mathbf{x})$ in latent space.

The framework optimises six competing objectives simultaneously:

$$\mathcal{L} = \mathcal{L}_{\text{recon}} + \alpha \cdot \mathcal{L}_{\text{KL}} + \beta \cdot \mathcal{L}_{\text{TC}} + \gamma \cdot \mathcal{L}_{\text{cluster}} + \delta \cdot \mathcal{L}_{\text{temporal}} + \epsilon \cdot \mathcal{L}_{\text{polarity}} \tag{2}$$

where $\mathcal{L}_{\text{recon}}$ is reconstruction loss, $\mathcal{L}_{\text{KL}}$ is the KL divergence to the GMM prior, $\mathcal{L}_{\text{TC}}$ penalises total correlation to promote disentanglement [9], $\mathcal{L}_{\text{cluster}}$ encourages well-separated GMM components, $\mathcal{L}_{\text{temporal}}$ enforces coherent microstate transitions between consecutive time points, and $\mathcal{L}_{\text{polarity}}$ enforces polarity invariance. Gradient conflict between objectives is managed through Projecting Conflicting Gradients (PCGrad) [27]. A staged training curriculum initialises the model with reconstruction and KL terms before progressively introducing clustering, temporal, and disentanglement objectives. The explainability framework operates at three levels, with each technique mapped to its intended audience. *Latent traversals* systematically vary individual latent dimensions while holding others fixed, generating topographic sequences that visualise what each dimension encodes intended for neuroscientists investigating latent factor structure. *Feature attribution* via Integrated Gradients [23] and SHAP [24], implemented using Captum and the SHAP library respectively, identifies which electrodes most influence cluster assignments intended for both researchers and clinicians. *Concept alignment* via TCAV [17] tests whether learned latent dimensions correspond to predefined neurophysiological concepts such as alpha power

distribution and dipole orientation intended for neuroscientists. *Counterfactual explanations* generate minimal perturbations to input topographies that change their cluster assignment, revealing decision boundaries in clinically actionable terms intended primarily for clinicians.

4.3. Evaluation

Evaluation spans five dimensions using Python-based tools including Scikit-learn and MNE-Python: (1) *Reconstruction quality* via Global Explained Variance (GEV); (2) *Disentanglement* via DCI [12], MIG [9], and SAP [13] metrics, adapted for EEG by defining ground-truth factors as anterior-posterior gradient, laterality, global field power, and dominant frequency band; (3) *Clustering validity* via silhouette scores, density-based spatial clustering of applications with noise (DBSCAN), calinski–harabasz index (CHI), and GMM component quality; (4) *Temporal fidelity* via transition matrix correlation with established microstate dynamics and microstate duration distribution matching; (5) *Neuroscientific validity* via canonical map recovery and source localisation consistency using standardised low-resolution electromagnetic tomography (sLORETA). Cross-dataset validation on the HBN dataset [28] tests generalisation across subjects.

5. Contributions to Date

The full experimental results addressing the overarching research question remain pending; substantial technical and scholarly groundwork has been established across four areas.

The Conv-VADE framework has been fully implemented in PyTorch, incorporating convolutional encoder and decoder components, a GMM-structured latent space, and the multi-objective loss function described in Section 3. The implementation currently produces latent representations and topographic reconstructions from EEG data and is undergoing systematic optimisation of hyperparameters and training curriculum. The LEMON dataset [25] has been acquired and a complete preprocessing pipeline implemented in MNE-Python, including bandpass filtering, ICA-based artifact rejection, average re-referencing, and GFP peak extraction. A baseline modified k-means implementation has been developed for downstream comparison against the Conv-VADE framework. A late-breaking work paper titled *"Interpretable EEG Microstate Discovery via Variational Deep Embedding: A Systematic Architecture Search with Multi-Quadrant Evaluation"* has been submitted for review. This paper reports the systematic architecture search conducted to identify the optimal Conv-VADE configuration and introduces a multi-quadrant evaluation framework for assessing disentanglement, reconstruction, clustering, and interpretability simultaneously. Early investigations have surfaced three technical challenges that inform the next research phase: (1) sensitivity of clustering quality to preprocessing pipeline choices, including reference scheme and bandpass filter parameters; (2) tension between reconstruction fidelity and latent disentanglement when the β weighting is increased; and (3) computational cost of full temporal modelling across long EEG recordings. These challenges are being addressed through systematic preprocessing robustness testing, staged training curricula, and efficient batching strategies respectively. The core experimental work focus on learning recurring patterns from EEG spatial topographic maps and validating subject-to-group generalisation. It constitutes the primary next research phase, detailed in Section 5.

6. Next Research Steps and Expected Final Contributions

The reviewed literature reveals that no existing method simultaneously addresses spatial richness, temporal coherence, explainability, and subject-to-group generalisation within a single framework for EEG microstate analysis. This research addresses that gap through the Conv-VADE multi-objective framework, which learns disentangled representations directly from EEG spatial topographic maps, preserves temporal transition dynamics, and produces interpretable brain state signatures generalisable across subjects. The remaining research is structured into three phases.

Disentanglement and Reconstruction (substantially underway): completion of the Conv-VADE architecture and multi-objective optimisation framework, including PCGrad gradient conflict resolution and systematic hyperparameter optimisation, forming the foundation for all subsequent work. *Temporal Transitions and Recurring Pattern Discovery*: implementation of the temporal coherence loss and validation of subject-specific to group-level generalisation, with selective cross-domain transfers including the VAMP-score temporal objective [20], the Information Bottleneck framework [29], and VQ-VAE discrete representations [19]. *XAI and Clinical Validation*: development of the full XAI toolkit including latent traversals, feature attribution, concept alignment via TCAV, and counterfactual explanations, with cross-dataset validation on the HBN dataset where the timeline permits. The minimum viable contribution comprises the first two phases in full plus latent traversals and feature attribution from the third. The expected final contributions are: a Conv-VADE framework for joint disentangled representation learning and clustering of EEG microstates; a multi-objective optimisation methodology balancing reconstruction, disentanglement, clustering, temporal coherence, and biophysical invariances; an XAI toolkit comprising latent traversals, attribution methods, concept alignment, and counterfactual explanations; evaluation protocols bridging machine learning and neuroscientific validity criteria; and empirical insights into sub-state variation within canonical microstate classes. These contributions aim to advance XAI for neural data and support the adoption of deep generative models in clinical and cognitive neuroscience applications where transparency is paramount.

Acknowledgments

This research was conducted at the Artificial Intelligence and Cognitive Load Research Lab, University College Cork, and supported by the Insight SFI Research Centre for Data Analytics.

Declaration on Generative AI

The authors used Grammarly to help with grammar check, spelling correction, and stylistic clarity during the preparation of this manuscript.

References

[1] D. Lehmann, H. Ozaki, I. Pal, EEG alpha map series: brain micro-states by space-oriented adaptive segmentation, Electroencephalography and Clinical Neurophysiology 67 (1987) 271–288.

[2] C. M. Michel, T. Koenig, EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review, NeuroImage 180 (2018) 577–593.

[3] A. Khanna, A. Pascual-Leone, C. M. Michel, F. Farzan, Microstates in resting-state EEG: current status and future directions, Neuroscience & Biobehavioral Reviews 49 (2015) 105–113.

[4] R. D. Pascual-Marqui, C. M. Michel, D. Lehmann, Segmentation of brain electrical activity into microstates: model estimation and validation, IEEE Transactions on Biomedical Engineering 42 (1995) 658–665.

[5] F. von Wegner, E. Tagliazucchi, H. Laufs, Information-theoretical analysis of resting state EEG microstate sequences: non-Markovianity, non-stationarity and periodicities, NeuroImage 158 (2017) 99–111.

[6] D. P. Kingma, M. Welling, Auto-Encoding Variational Bayes, in: Proceedings of the International Conference on Learning Representations (ICLR), 2014.

[7] Z. Jiang, Y. Zheng, H. Tan, B. Tang, H. Zhou, Variational deep embedding: an unsupervised and generative approach to clustering, in: Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI), 2017, pp. 1965–1972.

[8] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, A. Lerchner, beta-VAE: Learning basic visual concepts with a constrained variational framework, in: Proceedings of the International Conference on Learning Representations (ICLR), 2017.

[9] T. Q. Chen, X. Li, R. B. Grosse, D. K. Duvenaud, Isolating sources of disentanglement in variational autoencoders, in: Advances in Neural Information Processing Systems (NeurIPS), volume 31, 2018, pp. 2615–2625.
[10] M. M. Murray, D. Brunet, C. M. Michel, Topographic ERP analyses: a step-by-step tutorial review, Brain Topography 20 (2008) 249–264.
[11] D. Vidaurre, S. M. Smith, M. W. Woolrich, Brain network dynamics are hierarchically organized in time, Proceedings of the National Academy of Sciences 114 (2017) 12827–12832.
[12] C. Eastwood, C. K. I. Williams, A framework for the quantitative evaluation of disentangled representations, in: Proceedings of the International Conference on Learning Representations (ICLR), 2018.
[13] A. Kumar, P. Sattigeri, A. Balakrishnan, Variational inference of disentangled latent concepts from unlabeled observations, in: Proceedings of the International Conference on Learning Representations (ICLR), 2018.
[14] A. Theissler, F. Spinnato, U. Schlegel, R. Guidotti, Explainable AI for time series classification: a review, taxonomy and research directions, IEEE Access 10 (2022) 100700–100724.
[15] M. Guillemé, V. Masson, L. Rozé, A. Termier, Agnostic local explanation for time series classification, in: Proceedings of the IEEE International Conference on Tools with Artificial Intelligence (ICTAI), 2019, pp. 432–439.
[16] U. Schlegel, H. Arnout, M. El-Assady, D. Oelke, D. A. Keim, Towards a rigorous evaluation of XAI methods for time series, in: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 2019, pp. 4197–4201.
[17] B. Kim, M. Wattenberg, J. Gilmer, C. Cai, J. Wexler, F. Viegas, R. Sayres, Interpretability beyond feature attribution: quantitative testing with concept activation vectors (TCAV), in: Proceedings of the International Conference on Machine Learning (ICML), volume 80, 2018, pp. 2668–2677.
[18] A. Baevski, W.-N. Hsu, A. Conneau, M. Auli, Unsupervised speech recognition, Advances in Neural Information Processing Systems (NeurIPS) 34 (2021) 27826–27839.
[19] A. van den Oord, O. Vinyals, K. Kavukcuoglu, Neural discrete representation learning, in: Advances in Neural Information Processing Systems (NeurIPS), volume 30, 2017, pp. 6306–6315.
[20] A. Mardt, L. Pasquali, H. Wu, F. Noé, VAMPnets for deep learning of molecular kinetics, Nature Communications 9 (2018) 5.
[21] A. Chattopadhyay, P. Hassanzadeh, S. Pasha, Predicting clustered weather patterns: a test case for applications of convolutional neural networks to spatio-temporal climate data, Scientific Reports 10 (2020) 1317.
[22] R. Lopez, J. Regier, M. B. Cole, M. I. Jordan, N. Yosef, Deep generative modeling for single-cell transcriptomics, Nature Methods 15 (2018) 1053–1058.
[23] M. Sundararajan, A. Taly, Q. Yan, Axiomatic attribution for deep networks, in: Proceedings of the International Conference on Machine Learning (ICML), volume 70, 2017, pp. 3319–3328.
[24] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Advances in Neural Information Processing Systems (NeurIPS), volume 30, 2017, pp. 4765–4774.
[25] A. Babayan, M. Erbey, D. Kumral, J. D. Reinelt, A. M. F. Reiter, J. Röbbig, H. L. Schaare, M. Uhlig, A. Anwander, P.-L. Bazin, et al., A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults, Scientific Data 6 (2019) 180308.
[26] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, M. S. Hämäläinen, MEG and EEG data analysis with MNE-Python, Frontiers in Neuroscience 7 (2013) 267.
[27] T. Yu, S. Kumar, A. Gupta, S. Levine, K. Hausman, C. Finn, Gradient surgery for multi-task learning, in: Advances in Neural Information Processing Systems (NeurIPS), volume 33, 2020, pp. 5824–5836.
[28] L. M. Alexander, J. Escalera, L. Ai, C. Andreotti, K. Febre, A. Mangone, N. Vega-Potler, N. Langer, et al., An open resource for transdiagnostic research in pediatric mental health and learning disorders, Scientific Data 4 (2017) 170181.
[29] N. Tishby, F. C. Pereira, W. Bialek, The information bottleneck method, in: Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing, 2000, pp. 368–377.