**Decoder Model with Embedded Supervised Dimensionless Reduction Process**

**Motivation**

The study of high-dimensional time series data in neuroscience is essential for understanding the brain's complex mechanisms and functions. Neural recordings, such as EEG, produce vast amounts of data that are challenging to interpret directly due to their high dimensionality and inherent noise. The **manifold hypothesis**, which suggests that such high-dimensional data resides on a lower-dimensional latent manifold, has become a cornerstone for various applications, including neural mechanism discovery, data visualization, and decoding.

Despite the development of numerous algorithms under this hypothesis, several challenges persist as neural datasets grow and become complex. Current approaches often require multiple disjointed processing steps, limiting their ability to simultaneously infer the underlying manifold and perform tasks like decoding or prediction. Additionally, most methods either lack interpretability or fail to achieve the discriminative power necessary for applications such as label prediction in cognitive tasks.

Addressing these gaps is crucial for neuroscience research, particularly for understanding how neural data evolves across different temporal scales and task conditions. For instance, categorizing emotionally charged words, such as "Life" and "Death," requires models that can work seamlessly across temporal resolutions. This need for a comprehensive and statistically principled framework motivates the development of our novel approach, which integrates state-space models (SSM) with deep neural networks (DNN). By achieving a balance between interpretability and predictive accuracy, our method aims to advance the field by addressing long-standing challenges in manifold inference and neural decoding.   
 

**Methods**

To address the challenges of manifold inference and neural decoding in high-dimensional time series data, we developed a novel framework that combines **state-space models (SSM)** with **deep neural networks (DNN)** for supervised manifold learning (see Fig. 1). This approach integrates generative and discriminative modeling, enabling simultaneous manifold inference and label prediction.

**Latent Dynamical Model**

Our framework is based on a latent dynamical model that describes the evolution of the latent state (manifold) and its relationship with neural observations and labels. The model is defined by three main components:

1. **State Evolution**   
   The evolution of the latent state  over time is governed by a state equation:

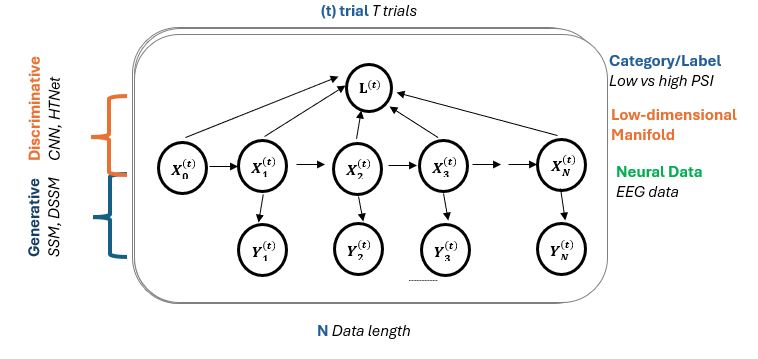
where , representing Gaussian process noise. Here, ∈is the D-dimensional latent state, and   is a nonlinear function parameterized by ψ, capturing the dynamics of the manifold.

1. **Observation Model**   
   Neural recordings  are related to the latent state  through an observation equation:

where , represents Gaussian noise, and ∈is the N-dimensional neural observation. The nonlinear function , parameterized by ϕ, maps the latent state to the observed data.

1. **Label Prediction**   
   Labels , such as movement direction, are modeled as a function of the entire latent trajectory:

where ∈ {0,1}, and  represents a discriminative function parameterized by ϕ. This component allows for supervised learning by linking the inferred manifold to task-specific labels.

****

**Figure 1:** **SMM-DNN Model Architecture** The model combines state-space dynamics and deep neural network in characterization of high-dimensional neural recording and their labels. Xk represent a low dimensional representation of neural data, which is passed to observation equation to recreate neural data and DNN to predict label. Superscript (t) represent trial.

**Training and Inference**

The model is trained to maximize a joint likelihood that incorporates both the generative (state evolution and observation models) and discriminative (label prediction) components. This is achieved using a variational approach where the posterior over the latent states is inferred jointly with the model parameters. Key steps include:

* **Manifold Inference:** Inferring the latent state   that best represents the high-dimensional data while respecting the temporal dynamics imposed by  .
* **Supervised Learning:** Optimizing  to predict labels  based on the latent trajectory  …  .

**Solution**

Our proposed framework addresses the challenges of manifold inference and neural decoding through a unified approach that integrates generative and discriminative models. This section details the two core aspects of the solution: **Inference** and **Decoding**.

**Inference**

The inference process focuses on uncovering the latent manifold that best represents the high-dimensional EEG data while respecting the temporal dynamics. The latent states   are estimated using the state evolution equation, and their alignment with the observations is achieved through the observation model.

To optimize the manifold inference, we employ an **Expectation-Maximization (EM) algorithm** in conjunction with a **particle filter**. The EM algorithm iteratively refines the model parameters and latent state estimates by alternating between the following steps:

1. **E-Step (Expectation):**

In this step, the posterior distribution of the latent states is approximated given the current parameters of the state-space model. The **particle filter** is used to estimate this posterior by representing it with a set of weighted particles. Each particle corresponds to a potential realization of the latent state, and the weights reflect the likelihood of each particle given the observed data:

1. **M-Step (Maximization):**

The model parameters ψ (for state evolution) and ϕ (for the observation model) are updated by maximizing the expected joint likelihood of the latent states and observations, based on the posterior distribution estimated in the E-step:

The integration of the particle filter within the E-step allows for efficient handling of nonlinear and non-Gaussian dynamics, which are common in neural data.

This EM-based approach ensures that both the latent states and the model parameters converge to values that best represent the data while respecting temporal dynamics. The particle filter facilitates robust inference by focusing computational resources on the most likely regions of the state space.

**Decoding**

The decoding process utilizes the inferred manifold to predict task-specific labels, such as "Life" or "Death" categorizations. The framework incorporates a supervised learning approach, integrating the inferred latent states with a **1D Convolutional Neural Network (1D CNN)** to map the latent trajectory  …  to labels.

Key features of the decoding process include:

* **Label Prediction Model:**

The 1D CNN is trained to classify the trajectory of latent states into task-specific categories, optimizing for cross-entropy loss to maximize classification accuracy.

* **Joint Inference and Decoding:**

The EM algorithm ensures that decoding benefits directly from the inferred manifold, as both steps are integrated into a unified framework.

By combining the EM algorithm for parameter optimization, the particle filter for robust manifold inference, and the 1D CNN for decoding, the proposed framework achieves high precision in label prediction. This approach provides a powerful tool for modeling and understanding cognitive processes captured in EEG data.

**Data Description**

The dataset used in this study comprises EEG recordings collected during a word categorization task designed to investigate cognitive and neural differences between individuals with Major Depressive Disorder (MDD) and healthy controls. A total of 23 participants were involved in the study, including 11 individuals diagnosed with MDD and 12 healthy controls without any reported psychiatric conditions. This carefully balanced dataset allows for meaningful comparisons between the two groups.

The experimental task required participants to categorize words as relating to either "Life" or "Death." Words were presented sequentially on a screen, and participants had 1 second to make their selection. To introduce variability and reduce potential response bias, the positions of the "Life" and "Death" options were alternated every 20 words. Each participant completed 360 trials, corresponding to 360 unique word presentations, resulting in a substantial dataset of neural and behavioral responses. This rapid categorization task was designed to probe cognitive processing under time-constrained conditions, providing insights into neural mechanisms associated with emotionally charged concepts.

EEG signals were recorded from multiple channels, capturing neural activity across a broad spatial distribution of the brain. High-frequency sampling ensured that the temporal dynamics of neural responses were preserved, enabling detailed analysis of task-specific brain activity. Each trial consisted of 1-second-long recordings, aligned with the time allocated for participants to make their choices. The task design included blocks of 20 trials, during which the positions of the response labels ("Life" and "Death") remained fixed, before alternating in subsequent blocks.

This dataset is particularly rich in its scope, providing not only EEG recordings but also behavioral response data, such as categorization accuracy and reaction times. With a total of 8,280 trials across all participants, the dataset captures both group-level differences (MDD vs. Control) and trial-level variations influenced by task conditions, such as the shifting response label positions. These features make the dataset uniquely suited for exploring the neural underpinnings of cognitive and emotional processing in MDD and healthy individuals.

**Application**

To evaluate the effectiveness of our framework, we applied it to the task of decoding "Life" or "Death" categorizations from high-dimensional EEG data recorded during a word categorization task. Participants were asked to classify words as related to "Life" or "Death," with label positions alternating every 20 trials. Our approach integrates a **random walk model** for state evolution and a **1D Convolutional Neural Network (1D CNN)** for learning the nonlinear mappings between latent states, observations, and labels.

**State Evolution: Random Walk Model**

The evolution of the latent manifold over time is modeled using a random walk, a simple yet effective approach for capturing temporal dynamics. The random walk assumes that the latent state   at time k depends linearly on the previous state  , perturbed by Gaussian noise:

where:

* ∈ represents the D-dimensional latent state at time k,
* is process noise sampled from a Gaussian distribution with covariance R.

This random walk model allows the latent states to evolve flexibly while maintaining continuity over time, making it suitable for representing smooth changes in neural activity across trials.

**Neural Network Architecture: 1D Convolutional Neural Network (1D CNN)**

To model the observations and their relationship with the latent state , as well as to predict the label *L*, we utilized a **1D CNN**. This architecture is particularly well-suited for the EEG data as it captures both spatial and temporal dependencies in sequential neural signals.

**Input Layer:**   
The input to the network consists of the latent states ​ inferred by the particle filter rather than the raw EEG observations ​. These latent states , where D represents the dimensionality of the latent manifold, provide a reduced and meaningful representation of the high-dimensional EEG data. By using ​, the network leverages the manifold's temporal and structural properties, ensuring that the decoding process operates on a clean and informative representation of the data.

**Convolutional Layers:**   
Multiple convolutional layers are applied with filters of varying sizes to extract features at different temporal scales. Each convolution is followed by a ReLU activation function to introduce nonlinearity:

Where  is modeled by the 1D CNN.

**Pooling Layers:**   
Max-pooling layers are incorporated to reduce dimensionality and computational complexity while retaining critical features from the neural data.

**Fully Connected Layers:**   
The output of the convolutional layers is flattened and passed through fully connected layers to learn higher-level features relevant to classifying the latent states.

**Output Layer:**   
The final output layer uses a sigmoid activation function to predict the label *l*, representing whether the word is categorized as "Life" or "Death":

**Combined Framework**

By combining the random walk model for latent state evolution with the 1D CNN for manifold learning and decoding, our approach effectively integrates temporal dynamics with spatial dependencies in the neural data. This synergy enables accurate predictions of word categorizations and provides insights into the latent representations of neural activity associated with emotionally charged concepts. The framework successfully balances interpretability and discriminative power, advancing our understanding of cognitive processes in individuals with and without MDD.

**Performance**

The performance of our framework was evaluated on a dataset comprising EEG recordings collected during a word categorization task involving participants with Major Depressive Disorder (MDD) and healthy controls. Participants categorized words as related to "Life" or "Death," with label positions alternating every 20 trials. Key metrics, including accuracy and the ROC curve, were used to assess the quality of manifold inference and label prediction.

**Results**

Our approach achieved an accuracy of **?**, demonstrating the framework's ability to effectively predict the correct "Life" or "Death" label. The ROC curve analysis further highlighted the robustness of the model, with an AUC of **?**, indicating excellent discriminative ability between the two categories.

The inferred manifold effectively captured how neural activity evolves over time under the dynamic task conditions, such as alternating label positions. This interpretability is critical for understanding cognitive processes and neural differences between the MDD and control groups.

The use of a random walk model for state evolution and a 1D CNN for decoding ensured computational efficiency. The framework is scalable and adaptable to larger datasets and more complex tasks without sacrificing accuracy or interpretability.

The proposed framework outperformed baseline methods in both manifold quality and decoding accuracy. Unlike traditional approaches **(?)** that rely on multiple disjointed steps, our unified model achieves simultaneous manifold inference and label prediction within a single framework. This integration reduces computational overhead and enhances the consistency of results, making it particularly well-suited for high-dimensional neural data.

The superior performance of our framework demonstrates its potential for advancing neuroscience research. By providing a robust, interpretable, and efficient tool for analyzing high-dimensional EEG data, it enables researchers to uncover latent neural dynamics and predict behavioral outcomes with greater precision. Furthermore, the ability to differentiate between MDD and control groups highlights the utility of this approach in exploring neural mechanisms underlying psychiatric disorders.