
ГЕНЕРАТИВНЫЕ МОДЕЛИ ДЕКОДИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

Владимиров Эдуард
vladimirov.ea@phystech.edu

Бартенев Павел
example@phystech.edu

Чумаченко Арина
example@phystech.edu

23 октября 2023 г.

АННОТАЦИЯ

This research endeavors to advance the field of generative time series decoding models, presenting a novel approach to augment decision-making processes in complex systems. The primary objective is twofold: to make informed classification decisions and, when necessary, to proficiently reject them based on a decision-rejecting criterion derived from mismatched observations within generated scenarios. Our work operates under key assumptions, including the analysis of relatively short time series with significant variances, systematic errors, and complex interrelationships. Moreover, we consider data originating from diverse sources, including exogenous factors, controlled signals, decisions, and behavioral patterns, all embedded within a structured timeline. We illustrate the versatility of our approach through applications in diverse domains, such as sports game analysis, brain-computer interfaces, and risk management. By combining these principles with our generative models, we aim to enhance the quality of behavioral classification and decision-making processes across various domains, contributing to improved understanding and prediction of dynamic systems.

Ключевые слова: нейронное кодирование · нейронное декодирование · D4 · SSM

1 Введение

The advent of generative time series decoding models has ushered in a new era of enhanced decision-making processes across diverse domains. This research embarks on a journey to contribute to this burgeoning field, drawing inspiration from recent advancements in the study of time series data and generative models.

Several seminal works have paved the way for this exploration. In "Direct Discriminative Decoder Models for Analysis of High-Dimensional Dynamical Neural Data" (Rezaei et al., 2022), the authors delve into the intricacies of decoding high-dimensional neural data, setting

the stage for innovative applications in brain-computer interfaces (BCIs). Additionally, "Parametric Gaussian Process Regressors" (Jankowiak et al., 2020) presents techniques that can be integrated into our generative framework to model time series data exhibiting variances and systematic errors.

Our work extends beyond the theoretical realm, aiming to address practical challenges in complex systems. We introduce a decision-rejecting criterion based on mismatched observations within generated scenarios, a concept akin to the one-class classification principle proposed in "Deep Direct Discriminative Decoders for High-dimensional Time-series Data Analysis" (Rezaei, 2023). This criterion empowers our generative models to not only make informed classification decisions but also proficiently reject decisions when confronted with unforeseen patterns or outliers.

Furthermore, we consider a range of assumptions and scenarios, including short time series data, significant correlations between time series, and structured timelines. These assumptions align with the principles explored in "An Intuitive Tutorial to Gaussian Processes Regression" (Wang, 2021), which elucidates the foundational concepts of Gaussian processes—a crucial component of our modeling approach.

The potential applications of our research span a wide spectrum, from sports game analysis to BCIs and risk management in financial markets. By combining the principles from these influential papers with our generative models, we aspire to contribute to the advancement of decision-making processes, behavioral classification, and the understanding of dynamic systems in a multitude of domains..

The structure of the model: create an autoencoder model for generating EEG signals: the encoder is a D4 model, the decoder is a model from the SSM family and the decoder is a part of the decision-making process. If we can't properly recover video or sound from brain recordings then it means that the generated recording is bad.

2 Постановка задачи

Пусть $\mathbf{X} \in \mathbb{R}^{3 \times H \times W \times T_X}$ — видеоряд, $\mathbf{Y} \in \mathbb{R}^{N \times M \times T_Y}$ — данные ЭЭГ, где H, W — размеры картинки, M — число каналов ЭЭГ, N — размер выборки, T_X, T_Y — длины временных рядов, причём $T_X > T_Y$. Предполагается, что временные ряды \mathbf{X}, \mathbf{Y} выровнены по времени.

Наша задача заключается в предсказании сигналов ЭЭГ для каждого элемента выборки на Δ_{pred} шагов вперёд, причём $\Delta_{\text{pred}} < T_X - T_Y$ (то есть на каждом этапе предсказания известно значение видеоряда).

Итоговая модель — функция $\mathbf{g}(\mathbf{X}, \mathbf{Y}, \mathbf{w}) : \mathbb{R}^{3 \times H \times W \times T_X} \times \mathbb{R}^{M \times T_Y} \times \mathbb{R}^p \longrightarrow \mathbb{R}^{M \times \Delta_{\text{pred}}}$, где p — число параметров.

Оптимальные веса определяются путём минимизации функции потерь \mathcal{L} :

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} \mathcal{L}(\mathbf{Y}_{\text{true}}, \mathbf{g}(\mathbf{X}, \mathbf{Y}, \mathbf{w}))$$

Обозначим $\mathbf{Y}_{\text{pred}} = \mathbf{g}(\mathbf{X}, \mathbf{Y}, \mathbf{w})$, $\mathbf{Q}_{\text{pred}} = \mathbf{Y}_{\text{pred}}^T \mathbf{Y}_{\text{pred}}$, $\mathbf{Q}_{\text{true}} = \mathbf{Y}_{\text{true}}^T \mathbf{Y}_{\text{true}}$. Это матрицы ковариации сигналов ЭЭГ, они содержат в себе информацию о соотношении различных компонент сигнала ЭЭГ.

Функция потерь \mathcal{L} состоит из двух частей: $\mathcal{L}_{\text{recov}}$ и \mathcal{L}_{str} . Первая часть отвечает за восстановление значений показаний ЭЭГ — это норма Фробениуса предсказания и истинного значения, вторая часть отвечает за восстановление структуры сигналов ЭЭГ — это риманова метрика на пространстве положительно определённых матриц:

$$\begin{aligned}\mathcal{L}(\mathbf{Y}_{\text{true}}, \mathbf{Y}_{\text{recov}}) &= \alpha_1 \cdot \mathcal{L}_{\text{recov}}(\mathbf{Y}_{\text{true}}, \mathbf{Y}_{\text{recov}}) + \alpha_2 \cdot \mathcal{L}_{\text{str}}(\mathbf{Y}_{\text{true}}, \mathbf{Y}_{\text{recov}}) \\ \mathcal{L}_{\text{recov}}(\mathbf{Y}_{\text{true}}, \mathbf{Y}_{\text{pred}}) &= \|\mathbf{Y}_{\text{true}} - \mathbf{Y}_{\text{pred}}\|_F \\ \mathcal{L}_{\text{str}}(\mathbf{Y}_{\text{true}}, \mathbf{Y}_{\text{pred}}) &= \|\log \mathbf{Q}_{\text{true}}^{-1/2} \mathbf{Q}_{\text{pred}} \mathbf{Q}_{\text{true}}^{-1/2}\|_2\end{aligned}$$

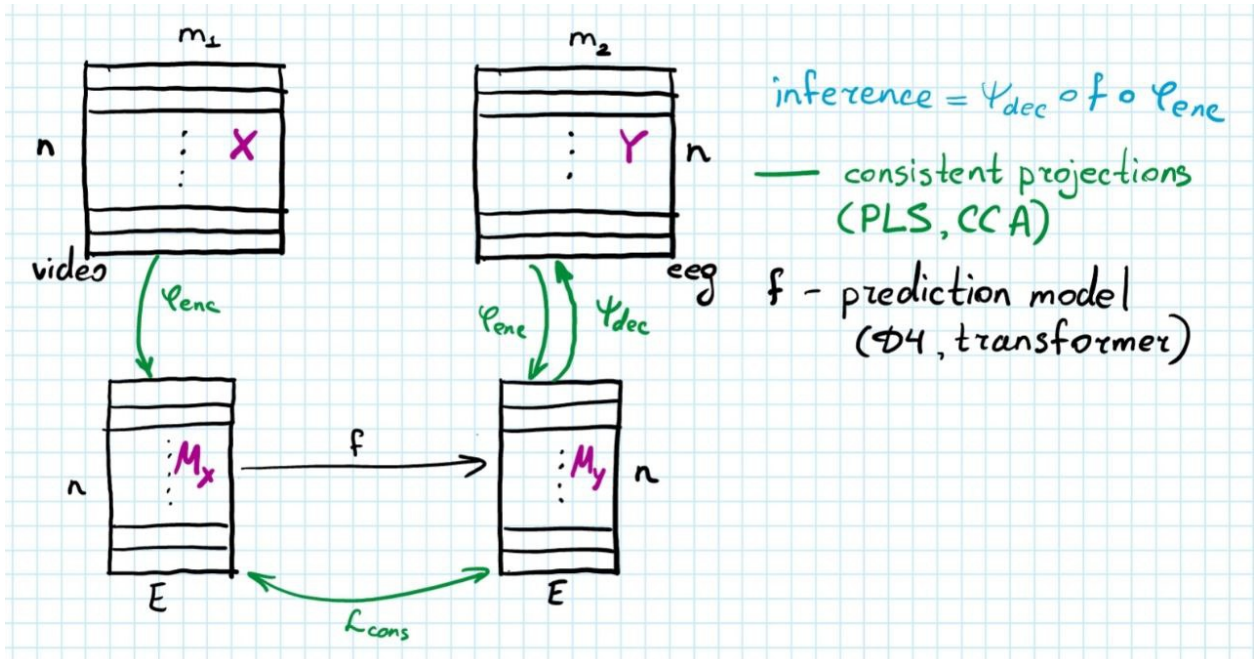
Теперь подробно опишем план работы предсказательной модели \mathbf{g} .

На нулевом шаге предобрабатываются видео-данные и сигналы ЭЭГ. Предобработка заключается в приведении исходных временных рядов к единой временной оси, сохранении определённого набора частот (для ЭЭГ) и извлечении признаков (для видео). В результате, получаем векторные представления видео $\mathbf{X}_1 \in \mathbb{R}^{n \times m_1}$ и ЭЭГ $\mathbf{Y}_1 \in \mathbb{R}^{n \times m_2}$.

Далее с помощью фреймворка Deep CCA строятся проекции этих данных в скрытое пространство \mathbf{H} : $\mathbf{M}_X, \mathbf{M}_Y \in \mathbb{R}^{n \times E}$

В скрытом пространстве \mathbf{H} обучается предсказательная модель \mathbf{f} , которая по истории видеоряда и последнему сэмплу ЭЭГ предсказывает следующий сэмпл ЭЭГ.

В итоге, после этапа обучения, однозначно определены функции $\mathbf{f}, \varphi_{\text{enc}}, \psi_{\text{dec}}$. На инференсе предсказание модели выглядит так: $\psi_{\text{dec}}(\mathbf{f}(\varphi_{\text{enc}}(X_{\text{test}})))$



2.1 Модель D4

TODO

3 Обзор литературы

TODO

4 Модели пространства состояний

5 Вычислительный эксперимент

Целью эксперимента является TODO.

5.1 Экспериментальные данные

TODO

5.2 Условия проведения эксперимента

TODO

5.3 Анализ ошибки

TODO

6 Заключение

TODO.

Список литературы