

Интерпретируемые пространства представлений для временных рядов

Владимир Александрович Богачев¹
Арчил Ивериевич Майсурадзе^{1,2}

¹Факультет ВМК, МГУ имени М.В. Ломоносова, Москва, Россия

²Институт ИИ, МГУ имени М.В. Ломоносова, Москва, Россия

21 декабря 2024

Вероятностная постановка задачи

Предполагается, что временные ряды $X_i \in \mathbb{R}^{T \times M}$ приходят из некоторого распределения $p(X \mid \theta_1, \dots, \theta_K)$.

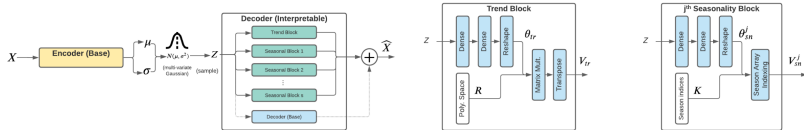
Задача построения интерпретируемых представлений заключается в построении отображения $\varphi(X)$ для которого $p(\varphi(X) \mid \theta_1, \dots, \theta_K)$ имеет заданный вид $p_0(\tilde{z} \mid \theta_1, \dots, \theta_K)$

Методы построения интерпретируемых представлений

Supervised: VAE-based, GAN-based

Unsupervised: VAE-based, SAE-based, topology-based

Архитектура TimeVAE



$$\mathcal{L}_{\text{VAE}} = \mathcal{L}_{\text{KL}} + \mathcal{L}_{\text{Rec}}$$

$$\mathcal{L}_{\text{KL}} = -\frac{1}{N} \sum_{n=1}^N \text{KL}(q_{\phi}(z | x_n) \parallel p(z))$$

$$\mathcal{L}_{\text{Rec}} = -\frac{1}{N} \sum_{n=1}^N \mathbb{E}_{q(z)} \log p_{\theta}(x_n | z)$$

Предложенная архитектура

Условный вариационный автокодировщик с дополнительным условием на независимость

$$\begin{aligned} p_{\phi}(z \mid x, \theta) &= \mathcal{N}(z \mid \mu(\theta), \Sigma(\theta)) = \\ &= \mathcal{N}(z_{J_0} \mid 0, I) \prod_k \mathcal{N}(z_{J_k} \mid \mu_{\phi_k}(\theta), \Sigma_{\phi_k}(\theta)) \end{aligned}$$

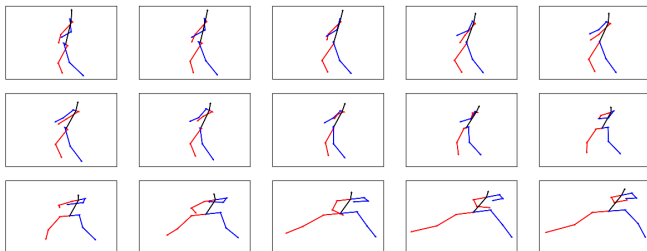
Сравнение моделей

	TV-TV	TV-PR	PR-TV	TV-TV
UPENN	0.054 ± 0.013	0.050 ± 0.012	0.052 ± 0.013	0.049 ± 0.01
Sine	0.073 ± 0.020	0.061 ± 0.018	0.064 ± 0.017	0.054 ± 0.01
Stockv	0.011 ± 0.012	0.010 ± 0.010	0.011 ± 0.010	0.010 ± 0.00

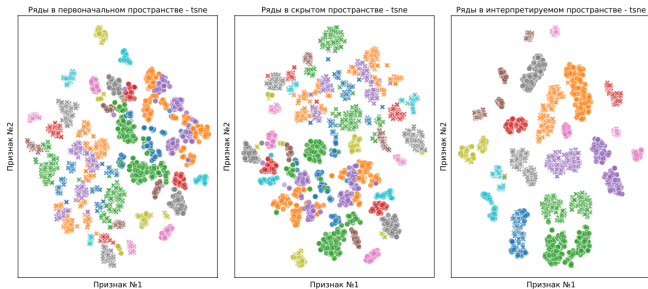
Сравнение моделей

	TimeVAE	Proposed
Disentanglement	0.21	0.91
Completeness	0.07	0.26
Informativeness	0.94	1.00

Решение задачи генерации



Применимость в дискриминативных задачах



Выводы

- Существуют supervised методы построения интерпретируемых представлений
- Перспективным продолжением является исследование unsupervised методов построения
- Намечен план исследования unsupervised методов