A theoretical framework for self-supervised contrastive learning for continuous dependent data

Alexander Marusov, Aleksandr Yugay, Alexey Zaytsev

fall into ML 2025
4th conference on machine learning & AI

Skoltech Applied AI
Applied Artificial Intelligence Center

Problem statement

Modern methods of self-supervised learning (SSL) allow to train neural networks without using labeled data, forming universal representations (embeddings) of the original signals.

Most of the existing self-learning methods are developed under the assumption of semantic independence of data, when objects are not interconnected in meaning.

However, for temporal and spatiotemporal data, this assumption is not fulfilled - neighbouring observations are correlated with each other. Modern approaches try to take into account semantics, but using heuristics. Therefore, it is necessary to develop a theory-based self-learning approach that takes into account the semantics of the data.

Current solutions

- Existing methods (TS2Vec, CoST, SoftCL) use the same loss functions as for images, assuming sample independence.
- SoftCL tries to account for correlations between time steps, but it does so heuristically, without rigorous theoretical justification.
- There is no single theoretical framework describing how the loss function should be designed for data with dependencies.
- Because of this, the models do not reflect the temporal or spatial structure of the data well and do not achieve optimal quality.

The proposed solution

We propose the DepTS2Vec method, a modification of TS2Vec with a dependency-aware loss function that takes into account time correlations. The method is based on the hypothesis of data continuity.

Two types of dependency between objects have been introduced:

- **hard** — only neighbouring points are similar:

$$g_{ij} = \begin{cases} 1, & \text{if } |i-j| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

- **soft** — the degree of similarity decreases with time distance:

$$a_{ij} = \exp\left(-\frac{(i-j)^2}{k}\right) \qquad g_{ij} = \begin{cases} \frac{a_{ij}}{\sum_{l=\min(i,j)+1}^{N} a_{\min(i,j)l}} & \text{if } i \neq j, \\ 0, & \text{otherwise.} \end{cases}$$

Loss function: $\mathcal{L}(\boldsymbol{\theta}) = -\sum_{i,j} g_{ij} \ln \widehat{g}_{ij}(f_{\boldsymbol{\theta}})$

where $\widehat{\mathbf{G}} = \{\widehat{g}_{ij}\}_{i,j=1}^{N}$ is the solution to the problem:

$$\begin{cases} \operatorname{Tr}(\mathbf{D}\widehat{\mathbf{G}}) + \mathcal{R}(\widehat{\mathbf{G}}) \to \min\limits_{\widehat{\mathbf{G}} \in \mathcal{G}_{\text{rsto}}}, \\ \text{s.t.} \quad \sum_{j=i+1}^{N} \widehat{g}_{ij} = 1 \quad \text{for any } i \in \overline{\{1, N\}}, \\ \sum_{i=j+1}^{N} \widehat{g}_{ij} = 1 \quad \text{for any } j \in \overline{\{1, N\}}. \end{cases}$$

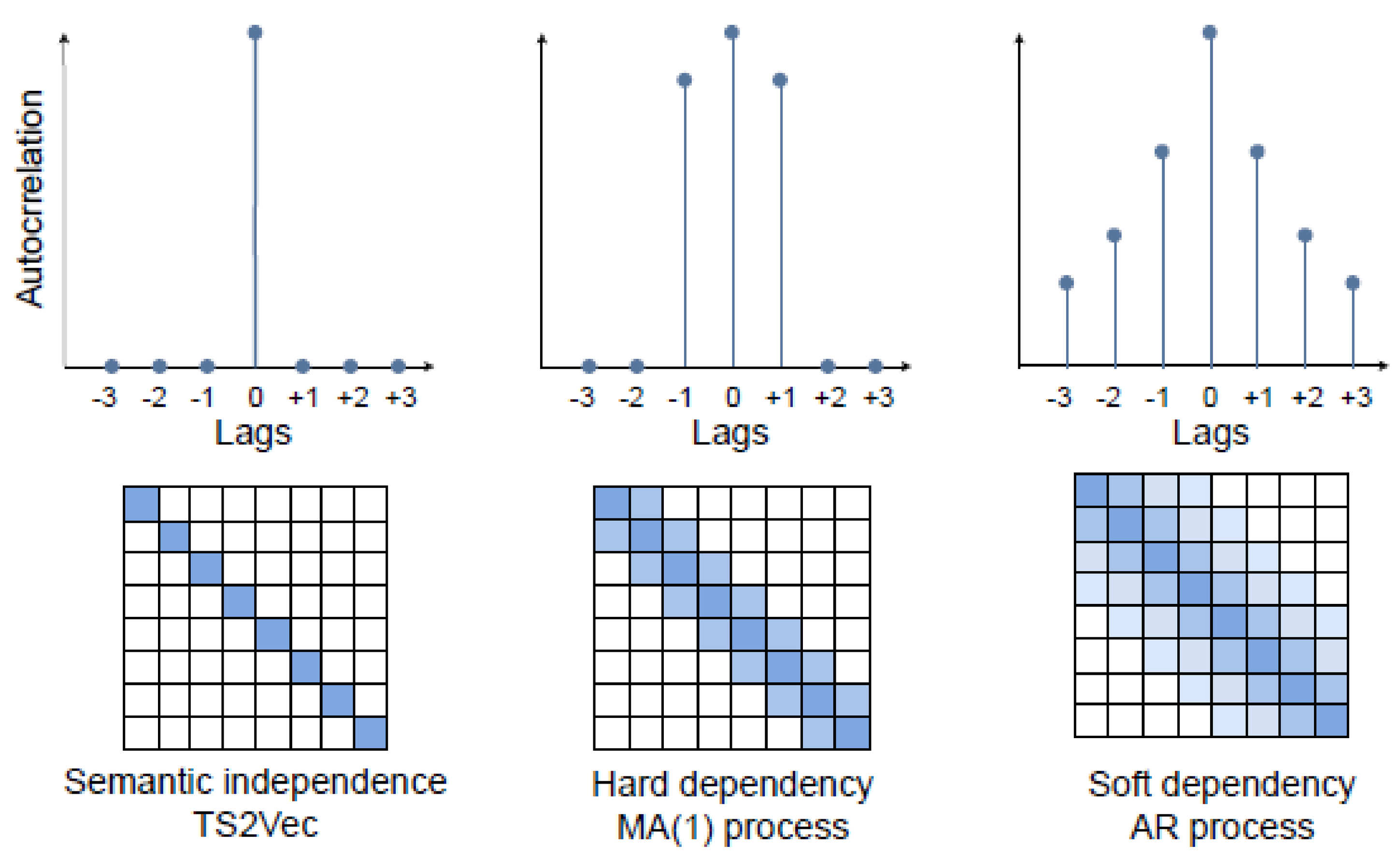

Fig. 1. Types of time dependence. We suggested considering the last two schemes (on the right).

Results

- On the UCR and UEA benchmarks, the DepTS2Vec method surpassed TS2Vec in accuracy by 2.08% and 4.17%, respectively (Table 1).
- Our method shows the best results under the assumption of data continuity (smooth dynamics, Table 2).
- In weather prediction tasks, our DepTS2Vec approach showed the best results in long-term drought forecasting (+7% ROC-AUC) and achieved the lowest RMSE when estimating temperatures for the week ahead (Table 3).

The results are applicable in climate modeling, time series analysis, and medicine, fields that use spatiotemporal data.

Model	UCR		UEA	
	Accuracy	Rank	Accuracy	Rank
TS2Vec [13]	82.56	2.35	68.72	2.58
SoftCL [16]	83.58	1.86	71.69	1.83
DepTS2Vec (Ours)	**84.64**	**1.79**	**72.89**	**1.58**

Table 1. Comparison of the average classification accuracy and the average rank on the UCR and UEA archive datasets.

	Type	# Datasets	TS2Vec	SoftCL	DepTS2Vec
UCR	HAR	21	2.21	**1.88**	1.90
	SENSOR	17	2.24	2.03	**1.74**
	DEVICE	12	2.33	2.08	**1.58**
	ECG	7	2.50	2.43	**1.07**
	MOTION	6	2.25	**1.83**	1.92
	TRAFFIC	2	3.00	**1.50**	**1.50**
UEA	HAR	9	2.22	2.11	**1.67**
	EEG	6	2.25	2.33	**1.42**
	MOTION	4	2.88	1.75	**1.38**
	ECG	2	3.00	**1.50**	**1.50**

Table 2. Average rank by data set type for UCR and UEA archives.

Model	Drought prediction (ROC-AUC ↑)					Temp. forecasting (RMSE ↓)
	1	2	3	4	5	
TS2Vec [13]	0.61	0.58	0.56	0.54	0.52	15.50
SoftCL [16]	0.67	0.64	0.60	**0.57**	0.55	10.33
DepTS2Vec (Ours)	**0.68**	**0.65**	**0.61**	**0.57**	**0.56**	**9.90**

Table 3. Comparison of methods for different space-time tasks.

Applied Artificial Intelligence Center
Contact: A.Marusov@skoltech.ru

Github

arXiv