

You're your own best teacher: A Self-Supervised Learning Approach For Expressive Representations in Time Series Generation

Johan Vik Mathisen

June 13, 2024

We are able to simultaneously reconstruct well and significantly improve downstream classification accuracy. Additionally, from the t-SNE plots ??, ?? and ??, we observe that NC-VQVAE can efficiently perform clustering as well, which is very interesting from a representation learning perspective. We improve both IS, FID and CAS for most datasets, indicating that the conditional distribution is better captured, as well as the synthetic data being closer to the ground truth.

NC-VQVAE is more adept at mimicking training data and less likely to overlook underrepresented classes. However, for datasets with limited training samples, our model can be prone to overfitting. Some challenges associated with TimeVQVAE remain, particularly in modeling data with sharp differences in modularity.

While issues persist, we believe our model represents a step forward. The representations learned by NC-VQVAE are more expressive than those from the naive VQVAE, showing that multiple objectives can be optimized without sacrificing reconstruction capability. These expressive representations facilitate capturing the semantics of the training data and producing synthetic samples with better global consistency.

In summary, NC-VQVAE demonstrate the ability to classify accurately and perform effective clustering while maintaining reconstruction capabilities, as well as enhancing generative quality.

0.1 Further work

There are many exiting directions this work can be taken further.

In [morningstar2024augmentations] they suggest that focus on augmentations is of great importance. The hunt for good augmentations in the time series domain is ongoing and should probably get more attention. An easy to use library of time series augmentations would be very beneficial to the community.

HF-LF split - augmentations tailored for HF and LF, as they often have quite different characteristics.

Further optimize the relationship between aug recon loss and choice of augmentations.

Improving on the stage 2 learning to better handle the expressive representations, and be able to create more diverse samples. Higher masking ratio during training, lower value for T etc.

The differences in Barlow and VibCReg indicate that further optimization of the SSL method/pretext task for generative performance is possible and could be

an interesting extension of this project.

Investigation of the attention maps. Could one find if there is a HF direction, translation direction etc. similarly to how one in NLP can find gender direction, nationality direction etc?

Develop better generative metrics.

More thorough analysis on the effect of the different losses.

Investigate the reason why the prior loss is not minimized, but we still observe much higher quality samples.

Understand the discrepancy between validation prior loss and generative quality. NC-VQVAE consistently produces samples with higher fidelity and improved generative scores, but it fails to model the latent space in such a way that the validation prior loss is minimized. The reason is a mystery, but would be very interesting to investigate further.

To better understand the geometry and topology of the discrete latent representations, and further investigate the effect of different augmentations and SSL methods, applying topological data analysis techniques is a possible direction of future research. To leverage