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

Johan Vik Mathisen

May 23, 2024

In this thesis we investigate possible improvements on the TimeVQVAE model presented by... We investigate how a "self supervised learning extension" of the tokenization affects the learned representations, and the effect on the prior learning. In particular we investigate if the learned representations are more informative, in the sense that they simultaneously enables high quality reconstruction, and improved the downstream classification accuracy. For the generative model we investigate if the learned representations enables faster convergence during training, and how the quality of the synthetic samples are affected.

0.1   Acknowledgements

Supervisors, Erlend, IDUN, UCR Archive creators

0.2   Motivation

- The role and importance of time series. - The need for models that capture complex structures for which traditional statistical models fail. Real world time series data is often incomplete (missing datapoints), irregular (datapoints not evenly spaced in time) and noisy. ML4ITS. - - Why do people care about time series generation (TSG)? - Applications - Why is (unsupervised) representation learning for time series interesting? Distributions of time series in their original temporal representation are complex and difficult to model. One would like to translate time series to a space where modelling is easier. This is one of the reasons to investigate representation learning for time series. Time series are recorded at record speed from sensors of various kinds (IoT, wearable devices). Unfortunately many of these do not have easily recognizable patterns for human observers, which makes labeling of such data quite difficult. In order to take advantage of this vast amount of unlabled data we need techniques that can extract useful patterns without supervision. This is one of the reasons for investigating possible unsupervised models. A subcategory of unsupervised learning called self-supervised learning has in recent times shown great potential for learning informative and useful representations without the need of labeled data in the fields of computer vision and natural language processing. Most notably the GPT models from OpenAI which utilizes masked language modelling for pre-training.

0.3   Overview/structure

- Main inspirations [**TimeVQVAE**]
- Collaboration with Erlend
- Structure of the thesis

0.4   Research questions

Stage1

RQ1: Does VQVAE learn good representations for classification?

RQ2: Will self a supervised learning approach enhance downstream classification while simultaneously reconstruct well?

RQ3: How does augmentations influence reconstruction and downstream classification?

Stage 2

RQ4: Will more expressive representations improve prior model learning?

RQ5: Does higher downstream classification accuracy correlate with improved class conditional generation?

RQ6: How does augmentations influence prior learning and synthetic sample quality?

- - - How does SSL VQVAE compare when we train a powerful prior model on top of it (MaskGIT)?

0.5   AI Declaration

- Ethical and environmental impact consideration with basis in UN sustainability goals