Proceedings of the Autumn Conference of
the Korean Data & Information Science Society
2024, 1–1

VAE-based Replication and Ensemble Methods for Enhanced Time Series Prediction†

Seung-Ho Ryu[1]· Young Eun Jeon[2]· Jung-In Seo[3]

[123]Department of Data Science, Andong National University

Abstract

This study proposes a novel approach to enhancing the predictive performance of time series data with seasonal variations, based on a variational autoencoder that incorporates seasonal patterns. To be specific, new time series data are generated using the variational autoencoder, and predictive modeling is conducted based on this generated data. For illustrative purposes, we utilize two real-world datasets: Air quality and sea ice extent. The procedure for applying the proposed method to the real datasets is as follows: First, missing values are imputed through a Kalman smoothing technique, and the time series data is decomposed into trend, seasonality, and remainder components. Second, the variational autoencoder model is trained by first transforming the sum of the seasonality and remainder components through variable transformation, and then processing them in the encoder. Seasonal patterns are incorporated into the model through the process of minimizing the Kullback–Leibler divergence between the seasonal latent vector from the seasonal prior distribution and the latent vector from the encoder. Finally, time series data without the trend is generated through the trained variational autoencoder, which is then back-transformed to the original scale, and the trend is added to create the replicated time series data. To evaluate the predictive performance of the proposed approach, a comparison between the original and replicated time series data is made using various predictive models such as recurrent neural networks and long short-term memory, and an ensemble technique. According to our results, the replicated time series data effectively captures the patterns of the original time series data. In addition, the predictive models trained using the replicated time series data outperform those trained using the original time series data. The proposed method can be applied to a wide range of time series data, with the potential to enhance predictive performance across various domains. Future research can extend this approach to accommodate multivariate and irregular time series, thereby broadening its applicability to even more complex datasets.

---

†This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2024S1A5A8026326).

[1] Graduate student, Department of Data Science, Andong National University, Andong, Korea.
[2] Researcher, Department of Data Science, Andong National University, Andong, Korea.
[3] Corresponding author: Associate professor, Department of Data Science, Andong National University, Andong, Korea. E-mail : leehoo1928@gmail.com

Keywords: Variational Autoencoder, Decomposition, Deep Learning, Time Series.