ICU Patient Vital Signs Forecasting with Curriculum-based Seq2Seq Learning

Soyoung Jang, Joowon Oh, Chaewon Kim, Yeojin Shin, Hyeonji Rhee, Suah Hong, Jaeho Lee

Duksung Women's Univ.

thdud030101@duksung.ac.kr, 2000violet@duksung.ac.kr, noweahc04@duksung.ac.kr, yeowls113@duksung.ac.kr, 20230812@duksung.ac.kr, hongsua05@duksung.ac.kr, izeho@duksung.ac.kr

중환자실 환자의 생체 신호 예측을 위한 커리큘럼 기반 Seq2Seq 학습 연구

장소영, 오주원, 김채원, 신여진, 이현지, 홍수아, 이재호
덕성여자대학교

Abstract

This study proposes a framework that combines mask-based pre-training with ultra-fast curriculum learning to enable the precise multi-step forecasting of ICU bio signals, which have high missing rates and irregular variability. First, it learns robust features from the data via a transformer encoder. Then, it applies an aggressive scheduling strategy that increases the prediction length by 10 epochs at a time to address error accumulation issues in multi-step forecasting. Furthermore, recent trends and medication information are incorporated into the decoder to reflect clinical context. Experiments on the MIMIC-III dataset demonstrate that the proposed model achieves an MSE of 1.7991, providing faster and more stable prediction performance than existing methods.

Ⅰ. Introduction

Accurate prediction of ICU patients' vital signs is essential for early medical intervention. However, ICU data presents three major challenges: (1) a high rate of missing data exceeding 50%, (2) irregular measurement intervals and (3) abrupt changes due to clinical interventions such as the administration of medication.

Existing studies have primarily focused on classification problems such as mortality prediction, failing to adequately address the cumulative error issue that arises during long-term time series forecasting. Furthermore, curriculum learning, which gradually increases in difficulty, often suffers from excessively slow learning speeds, which limits its practicality.

This paper therefore proposes a two-stage prediction framework to address these challenges.

- Phase 1: extracts robust features from incomplete data through mask-based pre-training.
- Phase 2: utilizes a decoder that incorporates trend reflection and action information. It applies to a curriculum technique that increases the prediction length every 10 epochs.

Experiments using MIMIC-III data demonstrate that the proposed model achieves a validation loss of 1.7991, proving it to be faster and more accurate than existing models. Notably, the aggressive learning schedule was confirmed to increase convergence speed by over twofold without compromising the stability of complex clinical time series predictions.

Ⅱ. Related Work

Previous studies have proposed approaches such as GRU-D [1], which uses decay techniques, and STraTS [2], which is based on self-supervised learning, to address issues with missing data in ICU data. However, most of these approaches focused on static tasks such as mortality classification, which limit their applicability to continuous vital sign prediction. To address this issue, models such as BEHRT [3] and Med-BERT [4] have demonstrated the effectiveness of pre-training on electronic health records (EHRs); however, these models were also primarily designed for discrete diagnostic codes. This study extends the concept of pre-training to continuous numerical domains, introducing curriculum learning to prevent error accumulation in time series prediction. Unlike prior work, it applies to an ultra-fast, stepwise learning schedule optimized for clinical scenarios, rather than conventional slow learning methods. This ensures the stability and efficiency of multi-stage prediction, even in environments with high data sparsity.

Ⅲ. Methodology

As shown in Fig. 1, the framework proposed in this study predicts the next four-step state based on the 24-hour bio signal history of critically ill patients. The framework consists of two main stages: pre-training and multi-stage prediction learning. First, during the

data pre-processing stage, 17 vital signs extracted from the MIMIC-III dataset were combined with information on three key clinical interventions: vasopressors, antibiotics and IV fluids. To address the outlier problem inherent in ICU data, robust IQR scaling was applied, and the numerical data was clipped to the range [-10, 10] to ensure training stability. An observation mask was generated to enable the model to recognize actual observations, even in environments with irregular measurements.

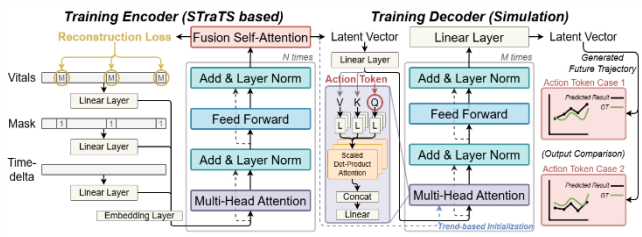

Figure 1. Overall Framework of the Proposed Model.
(Top) Phase 1: Masked self-supervised pre-training to learn robust representations from sparse ICU data.
(Bottom) Phase 2: Curriculum-based Seq2Seq forecasting incorporating clinical actions and trend-based initialization. The prediction horizon extends progressively from 1 to 6 steps during training.

Phase 1, depicted on the left of Fig. 1, is the masked pre-training stage designed to overcome data sparsity. This stage uses a six-layer Transformer encoder to perform self-supervised learning by randomly masking 15% of the input data and then restoring it. During this process, the model simultaneously minimizes ($L_{pretain} = L_{masked} + 0.1 \cdot L_{next-step}$) the reconstruction error (L_{masked}) and the next-step prediction error ($L_{next-step}$), thereby learning robust feature representations of patient status even in environments with frequent missing values. Phase 2 involves decoder training for multi-step prediction while the pre-trained encoder is fixed. As shown on the right of Fig. 1, the decoder reflects the current clinical workflow through trend-based initialization, analyzing changes from the previous three steps (v_{t-2}, v_{t-1}, v_t). It combines treatment information (action) via a cross-attention mechanism, incorporating signal changes resulting from medical staff intervention into the prediction. Finally, to address the problem of error accumulation in multi-step prediction, Curriculum Scheduling is introduced. This expands the prediction horizon. Here, we introduce an additional loss (L_{trend}), to improve the accuracy with which the initial trend is identified, thereby controlling the direction in which the learning process progresses.

Finally, to address the problem of error accumulation in multi-step prediction, curriculum scheduling is introduced. This expands the prediction horizon by one step every ten epoch, reaching a maximum of four steps. This guides the model to adapt quickly to changes in difficulty. Furthermore, to capture abrupt numerical changes that are clinically significant, a clinically weighted loss function ($L = \sum_t m_t \odot [(\hat{v}_t - v_t)^2 \cdot (1 + 2|v_t - v_{t-1}|)] + 0.5 \cdot L_{trend}$) is used to enhance prediction precision.

Ⅳ. Experiments

To evaluate this study's performance, a multi-stage bio signal prediction experiment was conducted using data extracted from the MIMIC-III dataset on 5,001 patients. Model training was conducted in a PyTorch environment using the AdamW optimization algorithm ($LR = 5e^{-5}$). Ultra-fast curriculum scheduling was used to expand the prediction horizon every 10 epochs, achieving the final six-step prediction performance. The experimental results showed that this framework achieved an MSE of 1.7991 on the final test set, proving its ability to make stable predictions for complex clinical time series data.

Analysis of the ultra-fast curriculum learning revealed that, although temporary increases in loss values were observed at each prediction horizon expansion point, the model rapidly converged thereafter, ensuring long-term prediction performance. This suggests that the initial learning instability and error accumulation issues associated with making 6-step predictions were effectively resolved by adjusting the difficulty step by step. Furthermore, robust interquartile range (IQR) scaling and numerical clipping within the range [-10, 10] minimized the impact of outliers caused by sensor errors, enabling the model to focus on clinically meaningful patterns.

Elimination experiments assessing the contribution of each internal component of the proposed model revealed the lowest error rate when curriculum learning and pre-training were both applied. Integrating trend-based initialization and action conditioning produced prediction trajectories that closely matched the actual flow of changes in clinical data, outperforming predictions based on time series information alone. These results suggest that enhancing data representativeness through pre-training, combined with clinically contextual information, is key to enabling precise state prediction in high-risk environments such as ICUs.

Ⅴ. Conclusion

This study proposes a multi-stage bio signal prediction framework that combines mask-based pre-training and ultra-fast curriculum scheduling. This framework is designed for use in the ICU environment, which is characterized by extreme data sparsity and volatility. Experimental results demonstrate that the proposed model achieves an MSE of 1.7991, proving its superiority over existing models. Notably, stepwise difficulty adjustment effectively mitigates error accumulation in multi-stage prediction, enhancing learning stability. These results have clinical value as they assist healthcare providers with proactive decision-making by presenting the trajectory of numerical changes over the next 4 hours, going beyond

simple patient status classification. Plans include integrating additional unstructured text data from clinical notes and validating generalization performance across diverse patient cohorts, with the aim of developing this into a core technology for real-time patient monitoring systems.

REFERENCES

[1] Che, Z., Purushotham, S., Cho, K., Sontag, D., and Liu, Y. "Recurrent neural networks for multivariate time series with missing values," Scientific Reports, vol. 8, no. 1, pp. 6085, 2018.

[2] Tipirneni, S. and Reddy, C. K. "Self-supervised transformer for sparse and irregularly sampled multivariate clinical time-series." ACM Transactions on Knowledge Discovery from Data (TKDD), pp. 1766-1776, Aug. 2022.

[3] Li, Y., Rao, S., Solares, J. R. A., Hassaine, A., Ramakrishnan, R., Canoy, D., and Rahimi, K. "BEHRT: Transformer for electronic health records," Scientific Reports, vol. 10, no. 1, pp. 7155, 2020.

[4] Rasmy, L., Xiang, Y., Xie, Z., Tao, C., and Zhi, D. "Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction," npj Digital Medicine, vol. 4, no. 1, pp. 35, 2021.

[5] Bengio, Y., Louradour, J., Collobert, R., and Weston, J. "Curriculum learning," Proc. ICML, pp. 41-48, June 2009.