Training Recurrent Neural Networks with Inherent Missing Data for Wearable Device Applications (Student Abstract)

Sutashu Tomonaga[1], **Haruo Mizutani**[2], **Kenji Doya**[1]

[1]Okinawa Institute of Science and Technology
[2]Suntory Global Innovation Center Limited
sutashu.tomonaga@oist.jp

Abstract

Wearable devices are transforming healthcare by providing continuous, real-time physiological data for monitoring and analysis. However, data often suffer from noise and significant missing values due to operational constraints and user compliance. Traditional approaches address these issues through data imputation during pre-processing, introducing biases and inaccuracies. We propose a novel method enabling Recurrent Neural Networks (RNNs) to inherently handle missing data without imputation. By implementing teacher-forcing during Backpropagation Through Time (BPTT) when data are available and switching to autonomous mode otherwise, our approach leverages RNNs' dynamics to model physiological signals accurately. We demonstrate our method's effectiveness using the Lorenz 63 system as a surrogate dataset, achieving robust reconstructions with 80% missing data.

Introduction

Wearable devices are integral to modern healthcare, continuously monitoring physiological signals like heart rate and temperature. They empower proactive health management and provide clinicians with real-time data.

Challenges with Wearable Device Data

Despite their benefits, wearable devices face inherent challenges:

- **Noise and Variability:** Uncontrolled environments introduce noise and high variability due to movement artifacts and environmental factors.
- **Missing Data:** Operational constraints like battery life and sensor connectivity issues cause irregular sampling and significant missing data.
- **Imputation Limitations:** Conventional pre-processing imputation can introduce biases and fail to capture true physiological dynamics.

Our Contribution

We propose a novel Recurrent Neural Network (RNN) training method that directly handles missing data, eliminating pre-processing imputation. Learning from incomplete data enhances robustness and accuracy in reconstructing physiological signals. We adapt *Scheduled Sampling* (Bengio et al. 2015) for scenarios with inherent missing data, making the model more applicable to real-world wearable datasets.

Background and Related Work

Dynamic Latent Variable Modeling

Recurrent Neural Networks (RNNs) are adept at modeling temporal dependencies and complex physiological dynamics (Hess et al. 2023), promising for estimating internal body states from wearable data.

Teacher-Forcing, Autonomous Mode, and Scheduled Sampling in RNN Training

In training RNNs, inputs can be provided in different ways. *Teacher-forcing* uses the true observed value as input for the current step, guiding the model with ground-truth information. Conversely, *autonomous mode* relies on the model's predicted outputs as inputs in subsequent steps (Doya and Selverston 1994). Autonomous mode can be (1) *with learning*, where parameters are updated based on output targets, or (2) *without learning*, where the model runs forward without updates.

Scheduled Sampling (Bengio et al. 2015) addresses the training-inference discrepancy by gradually shifting from teacher-forcing to autonomous with learning. It starts with 100% teacher-forcing and incrementally increases the autonomous mode probability. However, this assumes complete and fully observed training data, which is not true for wearable device data with large gaps and missing values.

To handle missing data, we introduce *Opportunistic Teacher-Forcing*, which switches between teacher-forcing and autonomous without learning based on data availability. Unlike scheduled sampling, our method does not use a predefined probability schedule but applies teacher-forcing opportunistically when data is present and switches to autonomous mode otherwise. By not requiring targets during autonomous steps, our approach naturally handles missing data from the start, overcoming scheduled sampling's limitations.

Proposed Method

Opportunistic Teacher-Forcing

Our method modifies the RNN training process to handle missing data without imputation. We introduce *Opportunistic Teacher-Forcing*, which switches between teacher-forcing and autonomous mode without learning based on data availability. This approach enables the RNN to learn dynamics even with significant data gaps, overcoming *Scheduled Sampling*'s limitations, which assumes complete data by starting with 100% teacher-forcing and gradually increasing autonomous mode probability.

Experiments and Results

Experimental Setup

- **Data Generation**: Simulate the Lorenz 63 system using standard parameters and numerical integration. The x-variable is observed, and the RNN is tasked with reconstructing x's dynamics from x alone.
- **Missing Data Simulation**: Introduce missing data by randomly omitting 80% of time steps.
- **Model Configuration**: Use a Multi-Timescale RNN (MT-RNN) with learnable time constants for hidden variables, inspired by (Doya and Selverston 1994).
- **Evaluation**: Assess performance based on reproducing the geometric attractor structure via delay embedding.
- **Training Details**: All models were trained using Backpropagation Through Time (BPTT) with identical early stopping criteria, concluding at 9,864 epochs for *Opportunistic Teacher-Forcing* and 11,046 epochs for *Scheduled Sampling*. For the latter, the autonomous mode probability was gradually increased from 0% to 80% in five equal increments.

Results

MT-RNN models trained with *Opportunistic Teacher-Forcing* successfully reconstruct the Lorenz system's dynamics comparable to those trained with *Scheduled Sampling*, despite 80% missing data. Figure 1 shows the true signal's delay embedding. Figures 2b and 2a illustrate that both training approaches yield robust autonomous reconstructions of the Lorenz attractor under 80% missing data.

Discussion

Implications for Wearable Device Data

Our findings indicate that RNNs can be effectively trained on datasets with significant missing data typical in wearable applications. Eliminating imputation enhances the model's ability to capture true physiological dynamics, leading to more accurate health monitoring.

Handling Noise and Variability

While focused on missing data, our method shows promise in handling inherent noise and variability by enabling the model to learn robust representations from data.

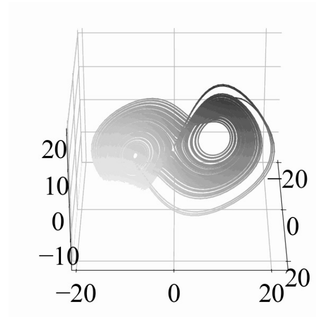
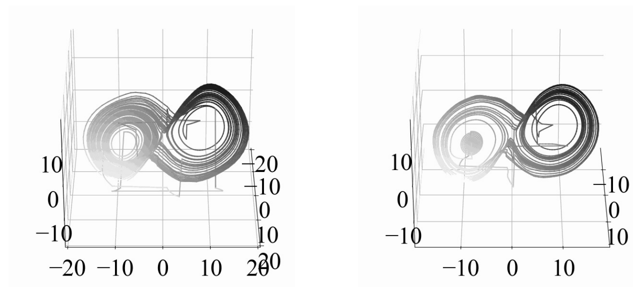


Figure 1: Delay embedding of true signal (x-variable of the Lorenz 63 system)



(a) Opportunistic Teacher-Forcing

(b) Scheduled Sampling

Figure 2: Delay embeddings of the reconstructed Lorenz attractor under autonomous mode for both *Opportunistic Teacher-Forcing* and *Scheduled Sampling*.

Limitations and Future Work

- **Testing on Real-World Data**: Our method has been evaluated only on surrogate data (the Lorenz 63 system). Validation on actual wearable device datasets is needed to confirm practical utility and robustness against real-world missing data patterns.
- **Missing Data Distribution**: We simulated missing data using a fixed probability, but real-world missing data may follow different distributions. Future research should investigate the method's performance under various missing data scenarios.

References

Bengio, S.; Vinyals, O.; Jaitly, N.; and Shazeer, N. 2015. Scheduled sampling for sequence prediction with recurrent neural networks. *Advances in neural information processing systems*, 28.

Doya, K.; and Selverston, A. I. 1994. Dimension Reduction of Biological Neuron Models by Artificial Neural Networks. *Neural computation*, 6(4): 696–717.

Hess, F.; Monfared, Z.; Brenner, M.; and Durstewitz, D. 2023. Generalized Teacher Forcing for Learning Chaotic Dynamics. *arXiv [cs.LG]*.