Neural Langevin-type Stochastic Differential Equations for Astronomical time series Classification under Irregular Observations

Paper

YongKyung Oh, Seungsu Kam, Dongyoung Lim, & Sungil Kim† | UNIST, Republic of Korea |
{yongkyungoh, lewki83, dlim, sungil.kim}@unist.ac.kr

Research Overview

- Addressing the classification challenges of irregular time series data in astronomical studies, this study leverages Neural Stochastic Differential Equations (Neural SDEs) to tackle data irregularity and incompleteness.
- We conduct a comprehensive analysis of the Neural Langevin-type SDEs' optimal initial conditions, which play a pivotal role in modeling the continuous latent state.
- Three different strategies for selecting the initial condition are compared under regular and irregular scenarios.

Proposed method

Neural Stochastic Differential Equations (Neural SDEs)

$$\boldsymbol{z}(t) = \boldsymbol{z}(0) + \int_0^t f(s, \boldsymbol{z}(s); \theta_f)\mathrm{d}s + \int_0^t g(s, \boldsymbol{z}(s); \theta_g)\mathrm{d}W(s)$$

where $\boldsymbol{z}(0) = h(\boldsymbol{x}; \theta_h)$, and $\{W(t)\}\_(t\geq 0)$ signifies a Brownian motion for the randomness in the process.

- $f(\cdot,\cdot;\theta_f)$ guides the systematic, predictable part of the motion.
- $g(\cdot,\cdot;\theta_g)$ accounts for the random fluctuations in the system.

Neural Langevin-type SDEs (Neural LSDEs)

$$\mathrm{d}\boldsymbol{z}(t) = \gamma(\bar{\boldsymbol{z}}(t); \theta_f)\mathrm{d}t + \sigma(s; \theta_\sigma)\mathrm{d}W(t)$$

where $\boldsymbol{z}(0) = h(\boldsymbol{x}; \theta_h)$, and the initial condition plays important role in evolving latent state.

$$\bar{\boldsymbol{z}}(t) = \zeta(t, \boldsymbol{z}(t), \boldsymbol{X}(t); \theta_\zeta)$$

where $\boldsymbol{X}(t)$ is the controlled path.

Initial condition selection

Because of the irregularity and missing data, we consider three different approaches to handle the initial condition using $\boldsymbol{x}$:

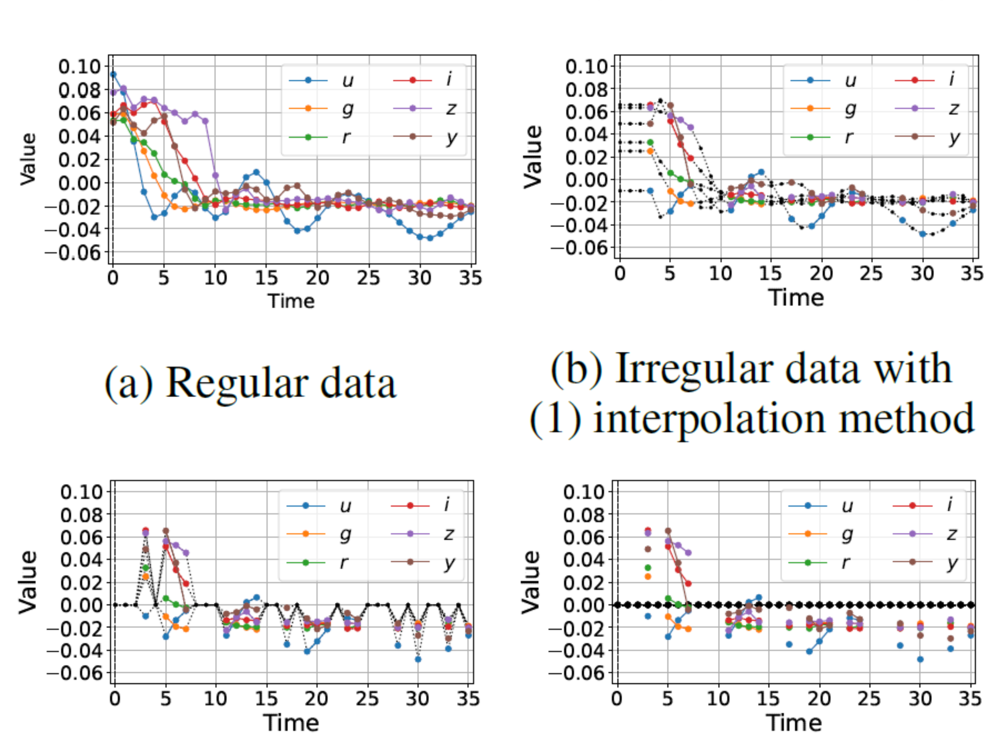
(1) Interpolation method: Apply natural cubic interpolation.

(2) Imputation method: Fill mean value for missing values.

(3) Static approach: Replace value of with zero.

We obtain $\boldsymbol{z}_0$ from $\boldsymbol{x}$ to determine the initial state $\boldsymbol{z}(0)$ at $t = 0$.

Figure 1. Example of regular and irregular (50% dropped) observation with the proposed three approaches



(a) Regular data

(b) Irregular data with (1) interpolation method

(c) Irregular data with (2) imputation method

(d) Irregular data with (3) static approach

$\boldsymbol{x} \in \mathbb{R}^{d_x \times T}$

↓

$\boldsymbol{z}_0 = \left(z_0^{(0)}, \ldots, z_0^{(T)}\right)$ with $z_0^{(0)} \in \mathbb{R}^{d_z}$

↓

$\boldsymbol{z}(0) = z_0^{(0)}$

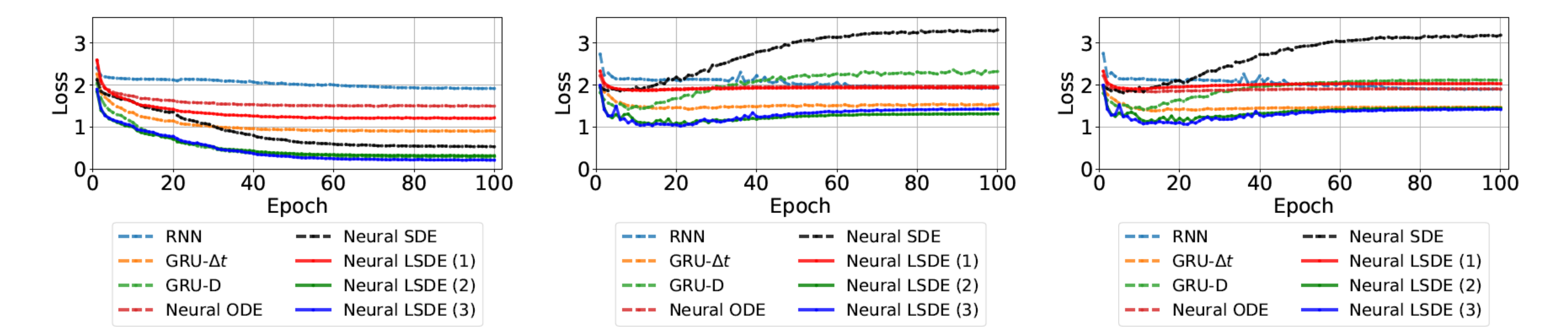
Experimental Results

LSST dataset refers to data from the Photometric Large Synoptic Survey Telescope (LSST) Astronomical Time Series Classification Challenge (PLAsTiCC).

✓ Regular setting vs. Irregular setting (Missing rate 50%).

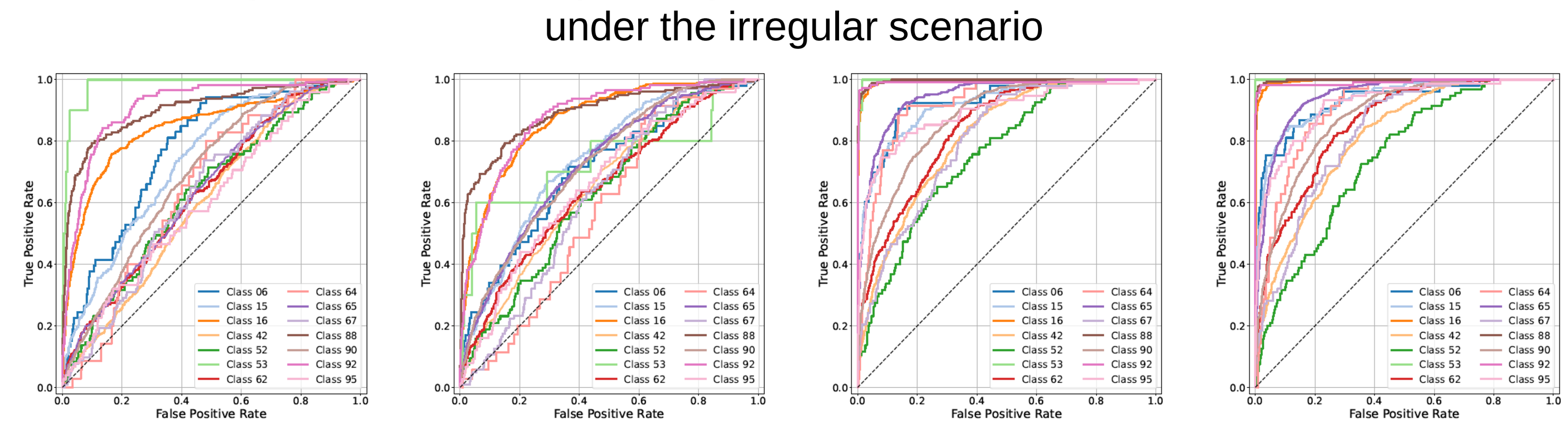
✓ 4925 samples, 6 input dimensions, 36 sequences, and 14 classes.

Figure 2. Comparing stability of loss with irregular setting using the selected methods



(a) Train loss (b) Validation loss (c) Test loss

Figure 3. Receiver operating characteristic curves for each class, under the irregular scenario



(a) Neural SDE (b) Neural LSDE (1) (c) Neural LSDE (2) (d) Neural LSDE (3)

Table 1. Classification performance on regular and irregular setting

Methods	Regular			Irregular		
	Accuracy	F1 score	AUROC	Accuracy	F1 score	AUROC
RNN	0.428 ± 0.054	0.218 ± 0.082	0.882 ± 0.032	0.344 ± 0.028	0.101 ± 0.031	0.819 ± 0.030
LSTM	0.524 ± 0.057	0.360 ± 0.057	0.919 ± 0.017	0.476 ± 0.024	0.316 ± 0.048	0.902 ± 0.010
BiLSTM	0.506 ± 0.032	0.327 ± 0.055	0.914 ± 0.008	0.445 ± 0.029	0.243 ± 0.036	0.890 ± 0.008
PLSTM	0.457 ± 0.030	0.273 ± 0.037	0.898 ± 0.006	0.426 ± 0.027	0.264 ± 0.047	0.876 ± 0.008
TLSTM	0.368 ± 0.077	0.139 ± 0.127	0.811 ± 0.052	0.332 ± 0.024	0.098 ± 0.056	0.809 ± 0.016
TGLSTM	0.491 ± 0.017	0.337 ± 0.013	0.912 ± 0.002	0.453 ± 0.023	0.261 ± 0.044	0.894 ± 0.010
GRU	0.604 ± 0.033	0.448 ± 0.039	0.947 ± 0.006	0.509 ± 0.046	0.355 ± 0.041	0.913 ± 0.016
GRU-Simple	0.354 ± 0.007	0.157 ± 0.026	0.824 ± 0.004	0.329 ± 0.005	0.086 ± 0.025	0.809 ± 0.007
GRU-Δt	0.540 ± 0.022	0.305 ± 0.026	0.927 ± 0.006	0.520 ± 0.023	0.300 ± 0.019	0.921 ± 0.004
GRU-D	0.551 ± 0.018	0.331 ± 0.039	0.929 ± 0.003	0.522 ± 0.022	0.327 ± 0.021	0.922 ± 0.004
Neural ODE	0.398 ± 0.014	0.153 ± 0.011	0.853 ± 0.004	0.394 ± 0.016	0.153 ± 0.017	0.850 ± 0.002
GRU-ODE	0.436 ± 0.054	0.230 ± 0.059	0.887 ± 0.018	0.434 ± 0.029	0.232 ± 0.053	0.887 ± 0.014
ODE-RNN	0.576 ± 0.021	0.381 ± 0.043	0.940 ± 0.005	0.542 ± 0.015	0.364 ± 0.023	0.929 ± 0.003
ODE-LSTM	0.412 ± 0.065	0.235 ± 0.107	0.850 ± 0.071	0.373 ± 0.059	0.164 ± 0.072	0.822 ± 0.053
Neural CDE	0.381 ± 0.009	0.161 ± 0.022	0.849 ± 0.003	0.372 ± 0.007	0.141 ± 0.022	0.845 ± 0.004
Neural RDE	0.317 ± 0.002	0.041 ± 0.011	0.796 ± 0.006	0.316 ± 0.001	0.037 ± 0.006	0.794 ± 0.003
Neural SDE	0.396 ± 0.016	0.210 ± 0.037	0.862 ± 0.005	0.390 ± 0.009	0.175 ± 0.010	0.856 ± 0.004
Neural LSDE (1)	0.402 ± 0.019	0.186 ± 0.019	0.866 ± 0.008	0.398 ± 0.031	0.183 ± 0.030	0.860 ± 0.009
Neural LSDE (2)	0.691 ± 0.012	0.556 ± 0.027	0.963 ± 0.002	0.638 ± 0.009	0.511 ± 0.018	0.953 ± 0.002
Neural LSDE (3)	**0.695 ± 0.009**	**0.573 ± 0.041**	**0.966 ± 0.001**	**0.648 ± 0.020**	**0.522 ± 0.027**	**0.956 ± 0.002**

Table 2. Ablation study of the model components in the proposed method

Methods			Regular			Irregular		
$\boldsymbol{z}(0)$	ζ	σ	Accuracy	F1 score	AUROC	Accuracy	F1 score	AUROC
(1)	X	L	0.403 ± 0.006	0.152 ± 0.018	0.861 ± 0.003	0.396 ± 0.004	0.152 ± 0.008	0.855 ± 0.006
		N	0.390 ± 0.015	0.172 ± 0.013	0.859 ± 0.003	0.391 ± 0.010	0.166 ± 0.008	0.852 ± 0.004
	O	L	0.412 ± 0.017	0.193 ± 0.027	0.868 ± 0.007	0.401 ± 0.019	0.172 ± 0.020	0.864 ± 0.006
		N	0.402 ± 0.019	0.186 ± 0.019	0.866 ± 0.008	0.398 ± 0.031	0.183 ± 0.030	0.860 ± 0.009
(2)	X	L	0.414 ± 0.015	0.212 ± 0.021	0.866 ± 0.006	0.349 ± 0.005	0.128 ± 0.032	0.822 ± 0.006
		N	0.428 ± 0.019	0.237 ± 0.052	0.869 ± 0.007	0.353 ± 0.018	0.136 ± 0.039	0.819 ± 0.014
	O	L	0.666 ± 0.015	0.534 ± 0.040	0.961 ± 0.002	0.638 ± 0.015	0.509 ± 0.030	0.954 ± 0.001
		N	0.691 ± 0.012	0.556 ± 0.027	0.963 ± 0.002	0.638 ± 0.009	0.511 ± 0.018	0.953 ± 0.002
(3)	X	L	0.315 ± 0.000	0.034 ± 0.000	0.791 ± 0.003	0.315 ± 0.000	0.034 ± 0.000	0.790 ± 0.002
		N	0.315 ± 0.000	0.034 ± 0.000	0.789 ± 0.002	0.315 ± 0.000	0.034 ± 0.000	0.790 ± 0.001
	O	L	0.685 ± 0.006	0.564 ± 0.018	0.964 ± 0.002	0.640 ± 0.012	0.524 ± 0.011	0.954 ± 0.002
		N	0.695 ± 0.009	0.573 ± 0.041	0.966 ± 0.001	0.648 ± 0.020	0.522 ± 0.027	0.956 ± 0.002

Reference

✓ Oh, Y., Lim, D., & Kim, S. (2024), Stable Neural Stochastic Differential Equations in Analyzing Irregular Time Series Data, The Twelfth International Conference on Learning Representations (ICLR) 2024, May 2024. (Spotlight)

Acknowledgement

This work was partly supported by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare, Republic of Korea (Grant number: HI19C1095), the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(No.RS-2023-00253002), the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(No.RS-2023-00218913), and the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2020-0-01336, Artificial Intelligence Graduate School Program (UNIST)).