Stable Neural Stochastic Differential Equations in Analyzing Irregular Time Series Data

Paper | Code | Slide | Author

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

TL; DR:

- We propose three stable classes of Neural Stochastic Differential Equations:
  - **Neural LSDE**: Neural Langevin-type SDE

$$dz(t) = \gamma(\bar{z}(t);\theta_t)dt + \sigma(t;\theta_\sigma)dW(t) \quad \text{with } z(0) = h(x;\theta_h).$$

  - **Neural LNSDE**: Neural Linear Noise SDE

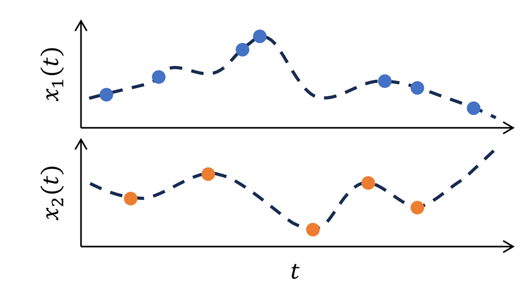
$$dz(t) = \gamma(t,\bar{z}(t);\theta_t)dt + \sigma(t;\theta_\sigma)z(t)dW(t) \quad \text{with } z(0) = h(x;\theta_h).$$

  - **Neural GSDE**: Neural Geometric SDE

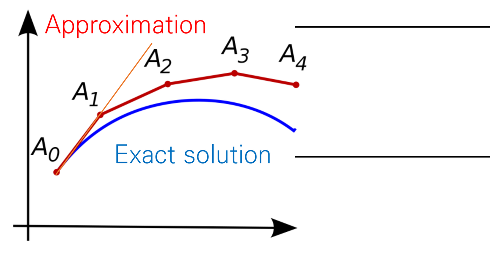
$$\frac{dz(t)}{z(t)} = \gamma(t,\bar{z}(t);\theta_t)dt + \sigma(t;\theta_\sigma)dW(t) \quad \text{with } z(0) = h(x;\theta_h).$$

- The proposed Neural SDEs are theoretically well-defined SDEs, with carefully designed drift and diffusion functions. Therefore, the stability of the proposed Neural SDEs, Neural LSDE, Neural LNSDE, and Neural GSDE, is guaranteed.
- Empirically, the proposed methods outperform existing techniques in a variety of tasks including interpolation, forecasting, and classification.

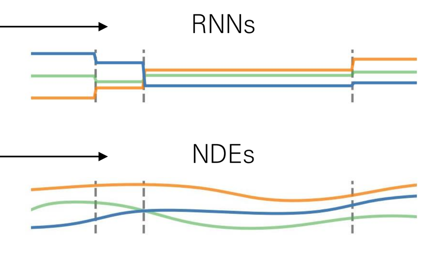
Challenges of "Irregular time series"



Examples of irregular time series



Find solution of differential equations
(Train RNNs vs. Solve NDEs)

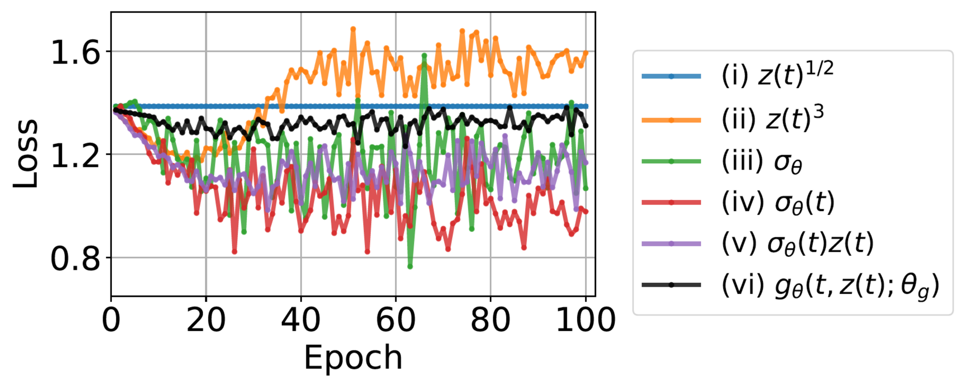


Hidden state trajectories
(Discrete vs. Continuous)

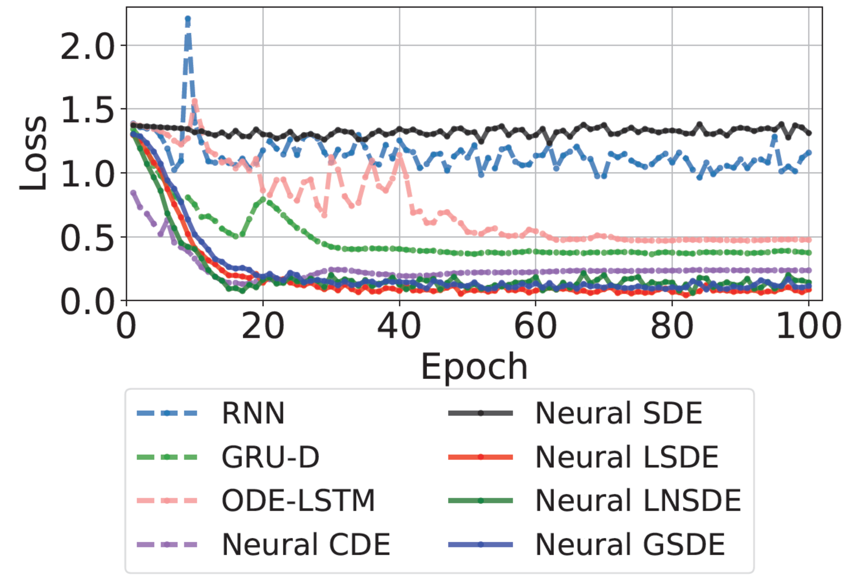
Variants of Neural Differential Equations.

- Neural ODEs: $z(t) = z(0) + \int_0^t f(s, z(s); \theta_f)ds$
- Neural CDEs: $z(t) = z(0) + \int_0^t f(s, z(s); \theta_f)dX(s)$
- Neural SDEs: $z(t) = z(0) + \int_0^t f(s, z(s); \theta_f)ds + \int_0^t g(s, z(s); \theta_f)dW(s)$

Proposed Stable Neural SDEs



Instability from stochastic noise



Impact of our proposed methods

Limitations of naïve Neural SDEs.

- Lack of guarantee of the existence and uniqueness of a strong solution.
- Stochastic destabilization, Instability with respect to Euler discretization.

Three classes of Neural SDEs.

- Building upon well-established SDEs, we propose three stable Neural SDEs:

Methods	Drift term	Diffusion term
Naïve Neural SDE (Neural SDE)	$f(t, z(t); \theta_f)$	$g(t, z(t); \theta_g)$
Neural Langevin-type SDE (Neural LSDE)	$\gamma(\bar{z}(t); \theta_\gamma)$	$\sigma(t; \theta_\sigma)$
Neural Linear Noise SDE (Neural LNSDE)	$\gamma(t, \bar{z}(t); \theta_\gamma)$	$\sigma(t; \theta_\sigma)z(t)$
Neural Geometric SDE (Neural GSDE)	$\gamma(t, \bar{z}(t); \theta_\gamma)z(t)$	$\sigma(t; \theta_\sigma)z(t)$

- We propose $\bar{z}(t)$ that incorporates a controlled path $X(t)$ in the drift term as follows:

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

where ζ is a neural network with parameter θ_ζ.

Key properties

- **Existence and Uniqueness**: The proposed Neural SDEs have their unique strong solutions.
- **Stochastic stability***: The stability of the proposed method is theoretically guaranteed.
- **Absorbing property of GSDE**: When a latent representation value hits zero, it remains zero.

* Stability in Neural Differential Equations ensures that the output distribution changes continuously with respect to input distribution shifts, guaranteeing the model's robust performance even when faced with irregular intervals or missing data.

Experiment Results

Interpolation

MSE error versus percent observed time points on PhysioNet Mortality

Methods	Mean Squared Error ($\times 10^{-3}$)				
RNN-VAE	13.418 ± 0.008	12.594 ± 0.004	11.887 ± 0.005	11.133 ± 0.007	11.470 ± 0.006
L-ODE-RNN	8.132 ± 0.020	8.140 ± 0.018	8.171 ± 0.030	8.143 ± 0.025	8.402 ± 0.022
L-ODE-ODE	6.721 ± 0.109	6.816 ± 0.045	6.798 ± 0.143	6.850 ± 0.066	7.142 ± 0.066
mTAND-Full	4.139 ± 0.029	4.018 ± 0.048	4.157 ± 0.053	4.410 ± 0.149	4.798 ± 0.036
LatentSDE	8.862 ± 0.036	8.864 ± 0.058	8.686 ± 0.122	8.716 ± 0.032	8.435 ± 0.077
Neural SDE	8.592 ± 0.055	8.591 ± 0.052	8.540 ± 0.051	8.318 ± 0.010	8.252 ± 0.023
Neural LSDE	**3.799 ± 0.055**	**3.584 ± 0.055**	3.457 ± 0.078	**3.262 ± 0.032**	**3.111 ± 0.076**
Neural LNSDE	3.808 ± 0.078	3.617 ± 0.129	**3.405 ± 0.089**	3.269 ± 0.057	3.154 ± 0.084
Neural GSDE	3.824 ± 0.088	3.667 ± 0.079	3.493 ± 0.024	3.287 ± 0.070	3.118 ± 0.065
Observed %	**50%**	**60%**	**70%**	**80%**	**90%**

Forecasting

MSE error performance versus percent observed time points on MuJoCo

Methods	Test MSE				Memory Usage (MB)
	Regular	30% dropped	50% dropped	70% dropped	
GRU-Δt	0.223 ± 0.020	0.198 ± 0.036	0.193 ± 0.015	0.196 ± 0.028	533
GRU-D	0.578 ± 0.042	0.608 ± 0.032	0.587 ± 0.039	0.579 ± 0.052	569
GRU-ODE	0.856 ± 0.016	0.857 ± 0.015	0.852 ± 0.015	0.861 ± 0.015	146
ODE-RNN	0.328 ± 0.225	0.274 ± 0.213	0.237 ± 0.110	0.267 ± 0.217	115
Latent-ODE	0.029 ± 0.011	0.056 ± 0.001	0.055 ± 0.004	0.058 ± 0.003	314
Augmented-ODE	0.055 ± 0.004	0.056 ± 0.004	0.057 ± 0.005	0.057 ± 0.005	286
ACE-NODE	0.039 ± 0.003	0.053 ± 0.007	0.053 ± 0.005	0.052 ± 0.006	423
NCDE	0.028 ± 0.002	0.027 ± 0.000	0.027 ± 0.001	0.026 ± 0.001	52
ANCDE	0.026 ± 0.001	0.025 ± 0.001	0.025 ± 0.001	0.024 ± 0.001	79
EXIT	0.026 ± 0.000	0.025 ± 0.004	0.026 ± 0.000	0.026 ± 0.001	127
LEAP	0.022 ± 0.002	0.022 ± 0.001	0.022 ± 0.002	0.022 ± 0.001	144
Neural SDE	0.028 ± 0.004	0.029 ± 0.001	0.029 ± 0.001	0.027 ± 0.000	234
Neural LSDE	0.013 ± 0.000	0.014 ± 0.001	0.014 ± 0.000	**0.013 ± 0.001**	249
Neural LNSDE	**0.012 ± 0.001**	0.014 ± 0.001	0.014 ± 0.001	0.014 ± 0.000	273
Neural GSDE	0.013 ± 0.001	**0.013 ± 0.001**	**0.013 ± 0.000**	0.014 ± 0.000	306

Classification

Accuracy and rank on 30 datasets with regular and three missing rates (30%, 50%, and 70%)

Methods	Regular datasets		Missing datasets (30%)		Missing datasets (50%)		Missing datasets (70%)	
	Accuracy	Rank	Accuracy	Rank	Accuracy	Rank	Accuracy	Rank
RNN	0.582 (0.064)	13.9	0.513 (0.087)	15.3	0.485 (0.088)	16.6	0.472 (0.072)	15.6
LSTM	0.633 (0.053)	11.2	0.595 (0.060)	12.1	0.567 (0.061)	12.9	0.558 (0.058)	12.5
GRU	0.672 (0.059)	8.3	0.621 (0.063)	10.2	0.610 (0.055)	10.3	0.597 (0.062)	10.3
GRU-Δt	0.641 (0.070)	10.3	0.636 (0.066)	8.9	0.634 (0.056)	8.7	0.618 (0.065)	10.2
GRU-D	0.648 (0.071)	10.3	0.624 (0.075)	10.4	0.611 (0.073)	11.1	0.604 (0.067)	10.8
MTAN	0.648 (0.080)	12.0	0.618 (0.099)	10.7	0.618 (0.091)	10.1	0.607 (0.078)	9.8
MIAM	0.623 (0.048)	11.0	0.603 (0.066)	11.1	0.589 (0.063)	12.2	0.569 (0.056)	12.3
GRU-ODE	0.671 (0.067)	9.8	0.663 (0.064)	9.5	0.666 (0.059)	8.3	0.655 (0.062)	7.8
ODE-RNN	0.658 (0.063)	9.1	0.635 (0.064)	9.3	0.636 (0.067)	8.2	0.630 (0.055)	8.5
ODE-LSTM	0.619 (0.063)	11.4	0.584 (0.064)	12.1	0.561 (0.065)	13.3	0.530 (0.085)	12.9
Neural CDE	0.709 (0.061)	8.2	0.706 (0.073)	6.4	0.696 (0.064)	6.5	0.665 (0.072)	7.6
Neural RDE	0.607 (0.071)	13.9	0.514 (0.064)	14.9	0.468 (0.068)	15.2	0.415 (0.077)	16.3
ANCDE	0.693 (0.067)	7.8	0.687 (0.068)	7.2	0.683 (0.078)	6.9	0.655 (0.067)	7.1
EXIT	0.636 (0.073)	11.1	0.633 (0.078)	10.2	0.616 (0.075)	10.6	0.599 (0.075)	11.1
LEAP	0.444 (0.068)	15.2	0.401 (0.078)	16.3	0.425 (0.073)	14.9	0.414 (0.070)	14.7
Latent SDE	0.456 (0.073)	16.6	0.455 (0.073)	15.4	0.455 (0.069)	15.0	0.446 (0.066)	15.1
Neural SDE	0.526 (0.068)	13.4	0.508 (0.066)	13.1	0.517 (0.058)	13.2	0.512 (0.066)	12.9
Neural LSDE	0.717 (0.056)	5.6	0.690 (0.050)	6.4	0.686 (0.051)	6.1	0.682 (0.067)	5.2
Neural LNSDE	**0.727 (0.047)**	**5.4**	**0.723 (0.050)**	**5.0**	**0.717 (0.054)**	**4.3**	**0.703 (0.054)**	**4.2**
Neural GSDE	0.716 (0.065)	5.7	0.707 (0.069)	5.3	0.698 (0.063)	6.1	0.689 (0.056)	5.3

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)).