Continuous CNN For Nonuniform Time Series

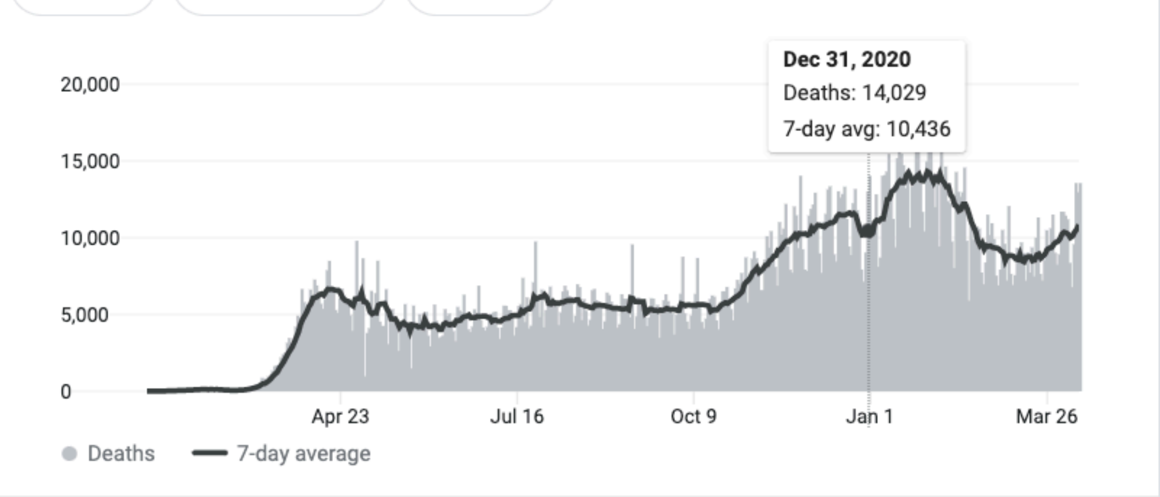
Hui Shi[1], Yang Zhang[2], Hao Wu[3], Shiyu Chang[2], Kaizhi Qian[3], Mark Hasegawa-Johnson[3], Jishen Zhao[1]

[1]UC San Diego [2] Watson AI Lab [3]University of Illinois at Urbana-Champaign

UC San Diego
MIT-IBM
ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

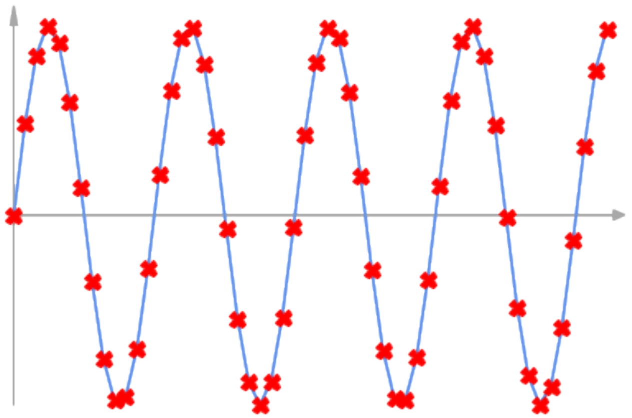
Time Series In The Real World

Uniform time series

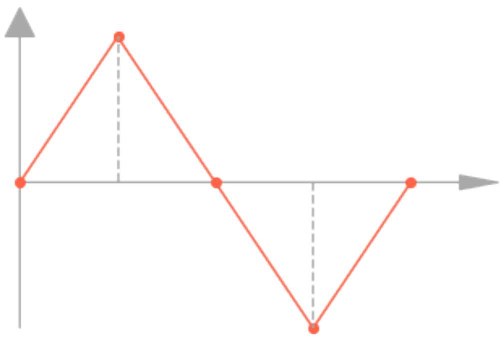


Nonuniform time series

Discrete Convolution On Time Series

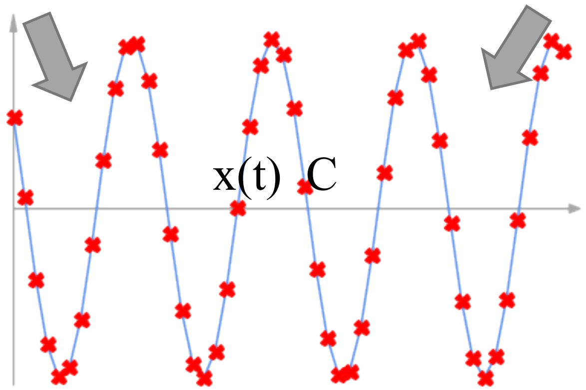


Signal x(t)

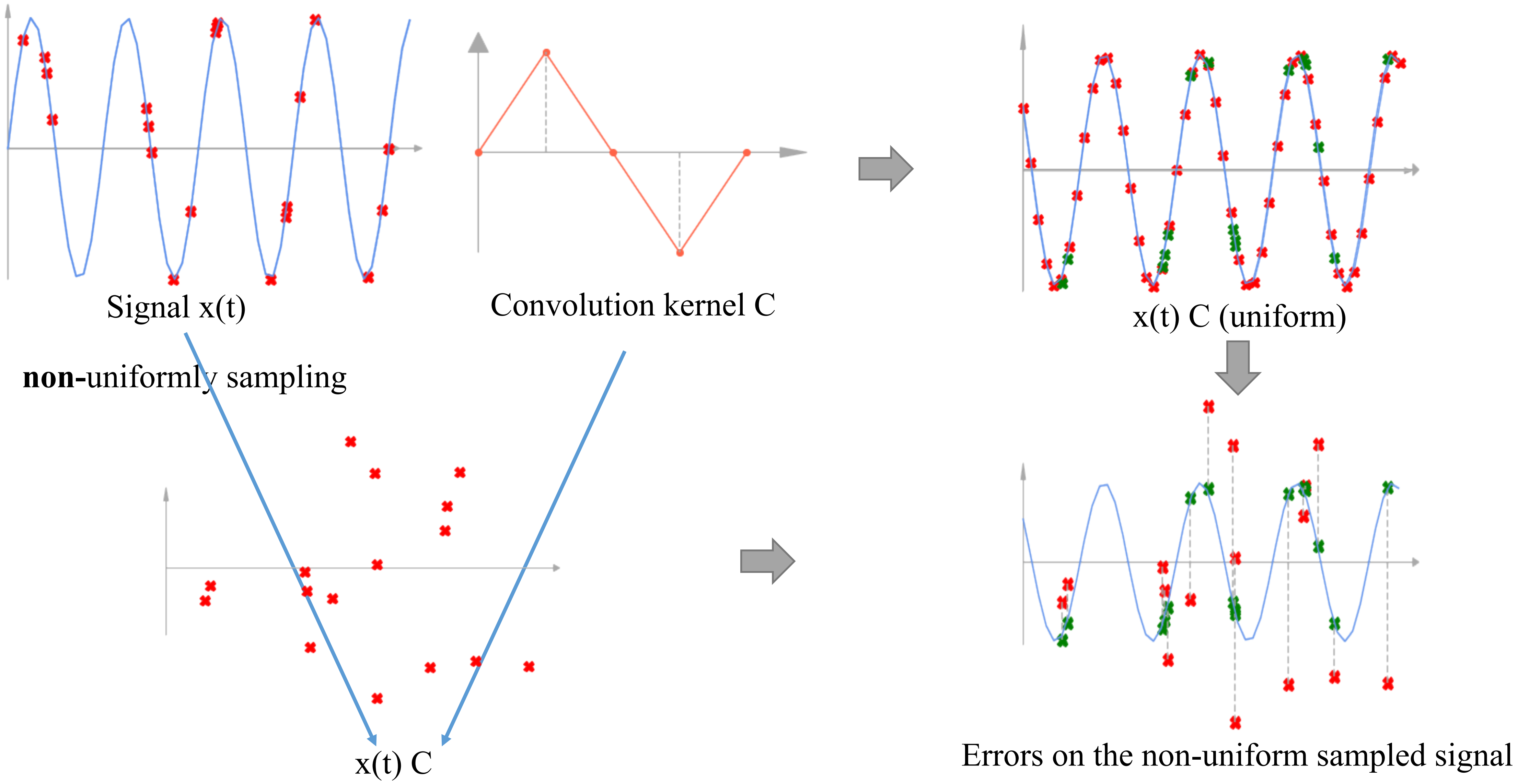


Convolution kernel C

uniformly sampling



Discrete Convolution On Time Series



Continuous CNN Theorem

Problem formulation

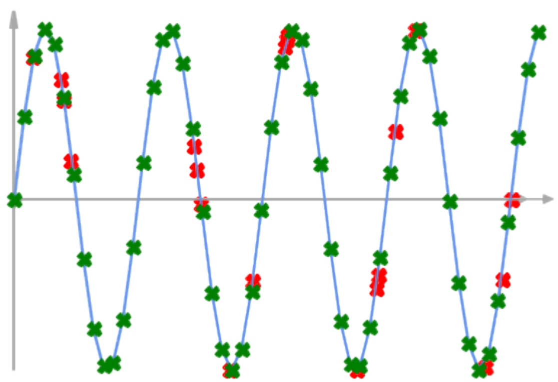
Observing signal $\mathcal{X}_{in}$ that is non-uniformly sampled from an unknown continues signal $x(\mathrm{t})$ with timestamps $\mathcal{T}_{in}$, the goal is to design a continuous convolutional layer that can produce output, $y(t_{out})$ for any arbitrary output time t_{out}

(1) Recover $\hat{\boldsymbol{x}}(\boldsymbol{t})$ via interpolation

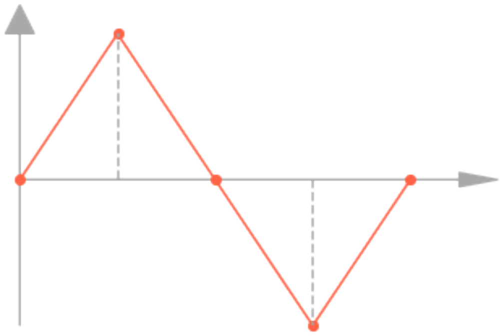
$$\hat{x}(t) = \Sigma_{i=1}^{N} x(t_i) I(t - t_i, \mathcal{T}_{in}, \mathcal{X}_{in}) + \epsilon(t, \mathcal{T}_{in}, \mathcal{X}_{in})$$

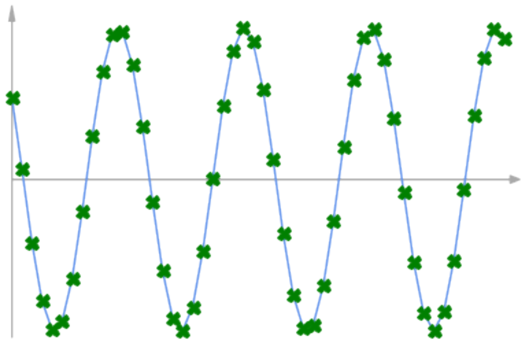
(2) Continuous convolution on $\hat{\boldsymbol{x}}(\boldsymbol{t})$

$$y(t_{out}) = [\hat{x}(t) * C(t)]|_{\mathrm{t=t_{out}}} + b$$


$\hat{\boldsymbol{x}}(\boldsymbol{t})$

$\mathcal{X}_{in}$


Convolution kernel C


$y(t_{out})$

Continuous CNN Theorem

Problem formulation

Observing signal $\mathcal{X}_{in}$ that is non-uniformly sampled from an unknown continues signal $x(\mathrm{t})$ with timestamps $\mathcal{T}_{in}$, the goal is to design a continuous convolutional layer that can produce output, $y(t_{out})$ for any arbitrary output time t_{out}

(1) Recover $\hat{x}(t)$ via interpolation

$$\hat{x}(t) = \Sigma_{i=1}^{N} x(t_i) I(t - t_i, \mathcal{T}_{in}, \mathcal{X}_{in}) + \epsilon(t, \mathcal{T}_{in}, \mathcal{X}_{in})$$

(2) Continuous convolution on $\hat{x}(t)$

$$y(t_{out}) = [\hat{x}(t) * C(t)]|_{\mathrm{t=t_{out}}} + b$$

$$y(t_{out}) = \Sigma_{i=1}^{N} x(t_i) [I(t - t_i, \mathcal{T}_{in}, \mathcal{X}_{in}) * \mathrm{C(t)}]|_{\mathrm{t=t_{out}}} + [\epsilon(t, \mathcal{T}_{in}, \mathcal{X}_{in}) * C(t) + b]|_{t=t_{out}}$$

$$= \Sigma_{i=1}^{N} x(t_i) K(t_{out} - t_i; \mathcal{T}_{in}, \mathcal{X}_{in}) + \beta(t_{out}, \mathcal{T}_{in}, \mathcal{X}_{in})$$

Continuous Convolution Layer

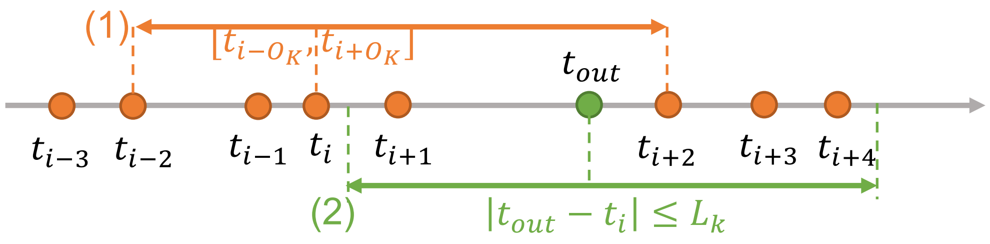
Assumptions:

(1) Stationarity and Finite Dependency

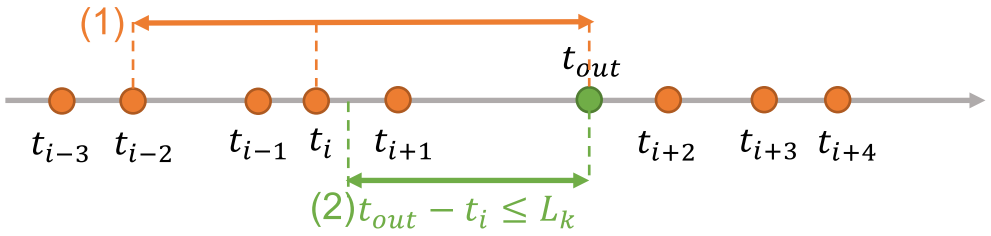
$K(t_{out} - t_i; \mathcal{T}_{in}, \mathcal{X}_{in}) = K(\{t_{out} - t_{i \pm o_K}, x(t_{i \pm o_K}\})$

(2) Finite Kernel Length

$K(t_{out} - t_i; \mathcal{T}_{in}) = 0, \forall |t_{out} - t_i| > L_k$



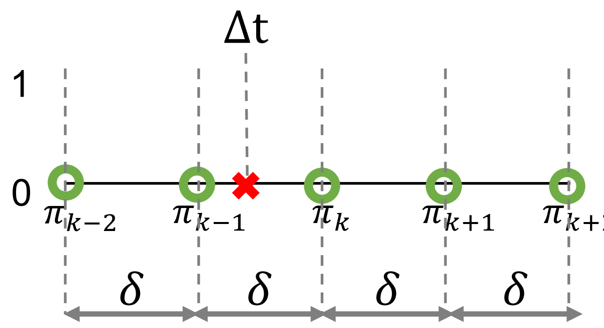
Causal setting for forecasting task:

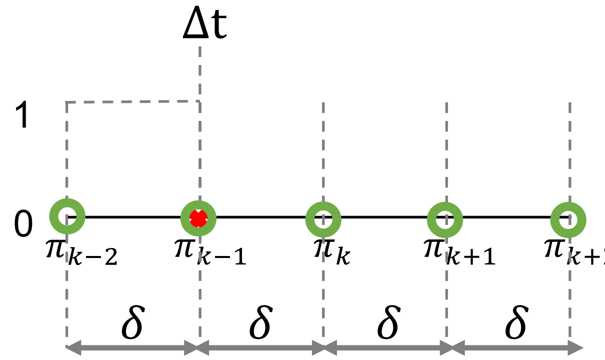


Two-hot encoding:

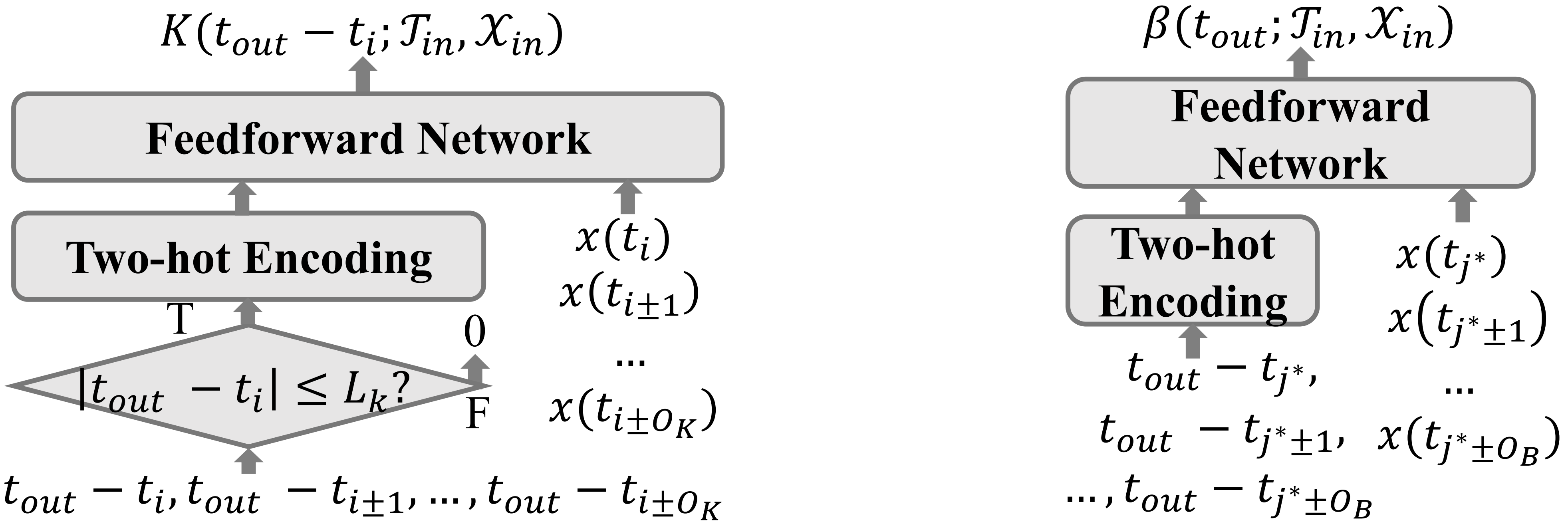
$$g_k = \frac{\pi_k - \Delta t}{\delta}; g_{k+1} = \frac{\Delta t - \pi_{k-1}}{\delta};$$

$$g_l = 0, \forall l \notin \{k-1, k\}$$



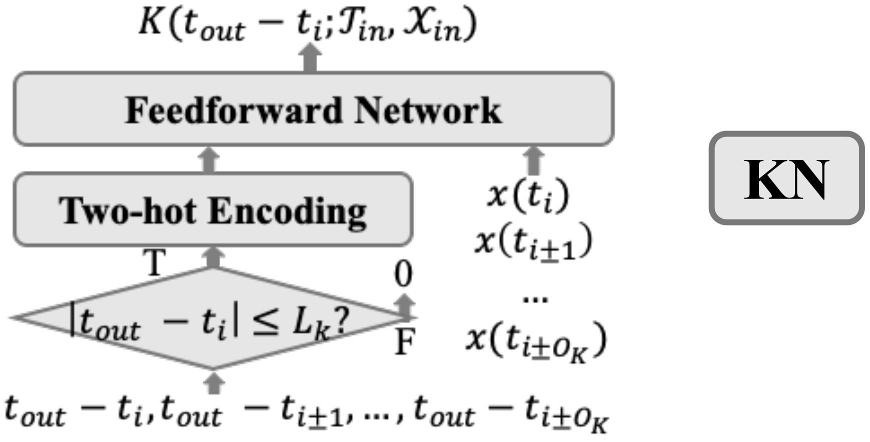


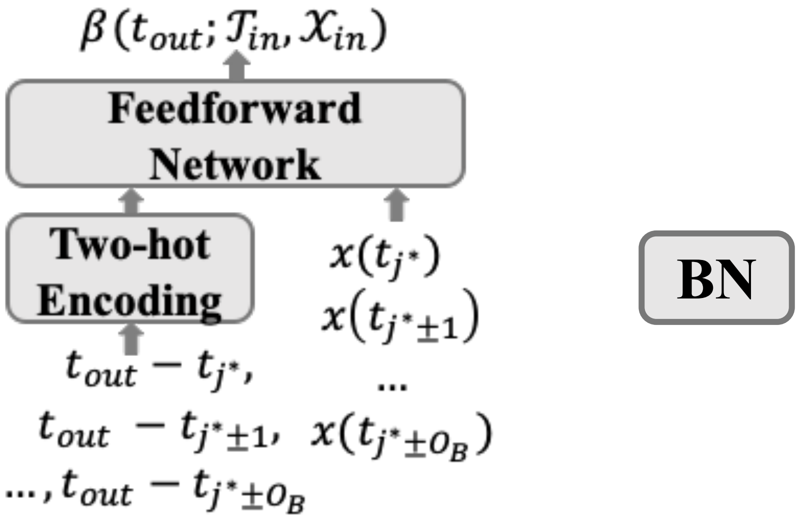
Continuous Convolution Layer

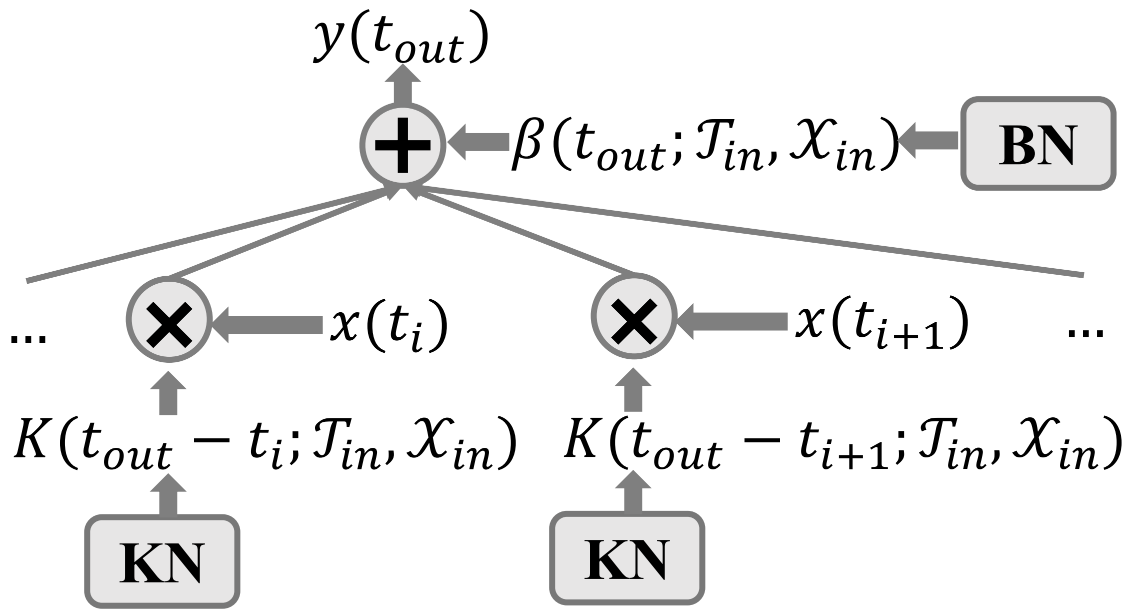


(a) Kernel network **(b) Bias network**

Continuous Convolution Layer

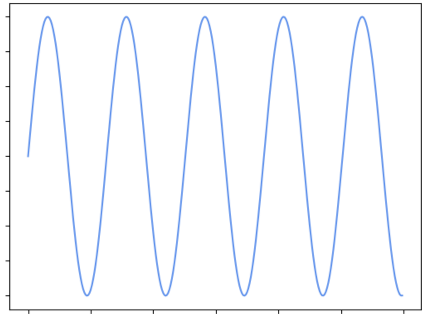






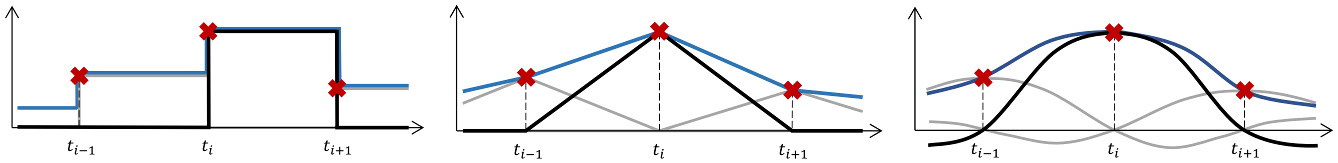
Continuous convolution layer

Evaluations: Autoregression

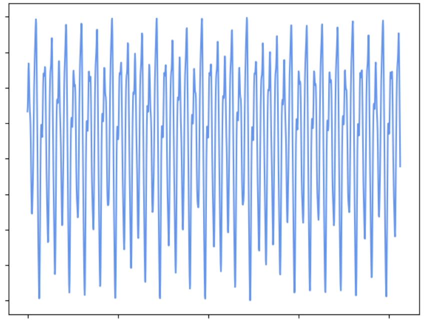


Sine:

$$x(t) = \sin(\frac{2\pi t}{T})$$

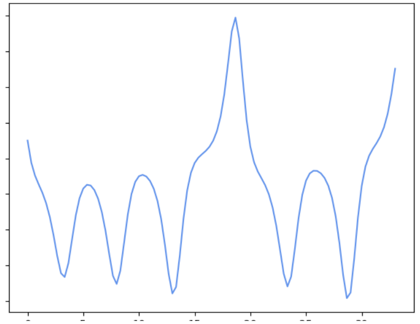


(a) Point-wise(-P) (b) Linear(-L) (c) Sinc(-S)



Mackey-Glass:

$$\dot{x}(t) = \beta \frac{x(t-\tau)}{1 + x(t-\tau)^n} - \gamma x(t)$$



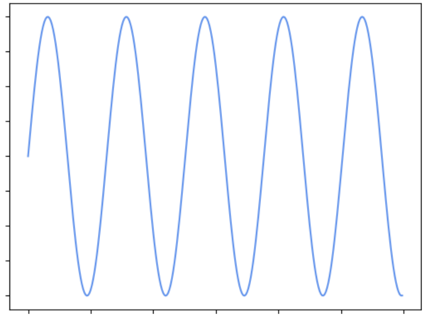
Lorenz:

$$\dot{x}(t) = \sigma\big(y(t) - x(t)\big)$$
$$\dot{y}(t) = -x(t)z(t) + \gamma x(t) - y(t)$$
$$\dot{z}(t) = x(t)y(t) - bz(t)$$

Alg.	**Sine**	**MG**	**Lorenz**
CNN	46.0 (8.22)	12.8 (3.92)	9.90 (3.33)
CNNT	20.2 (7.65)	3.50 (1.29)	5.97 (2.41)
CNNT-th	8.44 (4.58)	3.00 (1.21)	8.37 (3.24)
ICNN-L	1.13 (0.87)	0.97 (0.53)	5.81 (2.78)
ICNN-Q	0.75 (0.65)	0.83 (0.46)	5.08 (2.59)
ICNN-C	0.72 (0.83)	0.72 (0.42)	4.22 (2.27)
ICNN-P	20.5(6.43)	1.95(0.79)	8.50(3.32)
ICNN-S	17.2(5.57)	3.51(1.36)	8.20(3.31)
RNNT	36.1(12.9)	8.15(3.32)	13.4(3.95)
RNNT-th	19.5(6.48)	8.48(3.11)	13.9(4.36)
CCNN	0.88 (0.61)	2.46 (0.89)	3.93 (1.73)
CCNN-th	**0.42** (0.36)	**0.53** (0.97)	**3.25** (1.67)

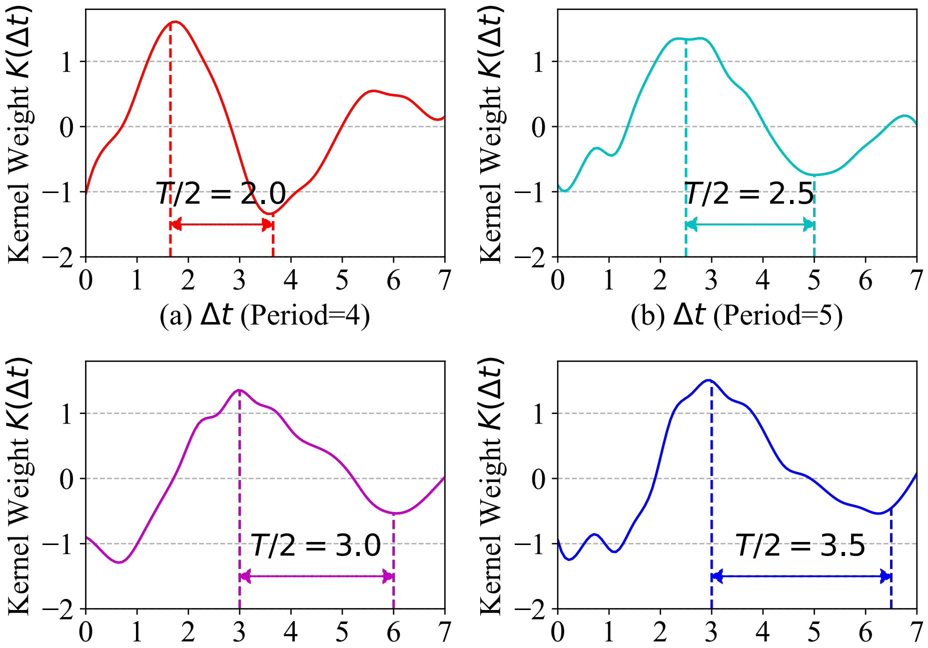
MSE loss in 10^{-2}

Evaluations: Autoregression



Sine:

$$x(t) = \sin(\frac{2\pi t}{T})$$

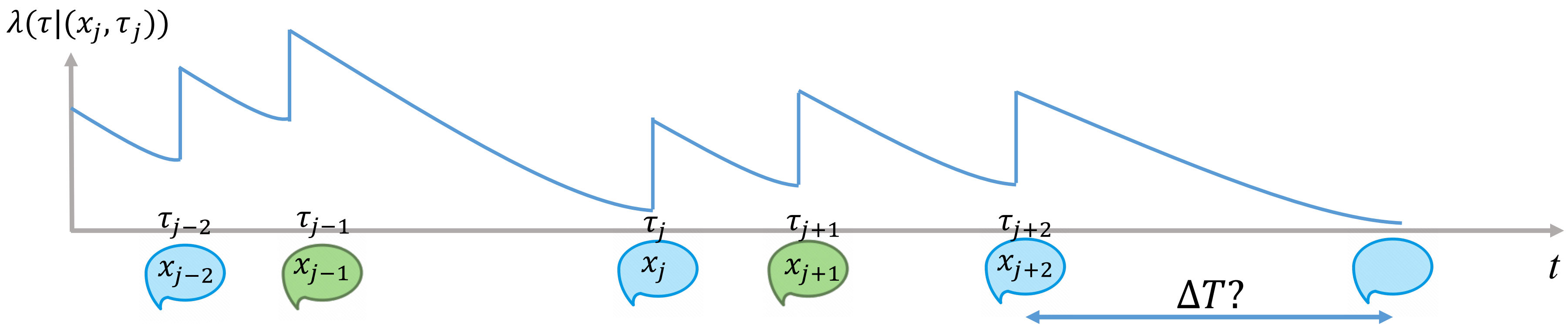


(a) Δt (Period=4) (b) Δt (Period=5) (c) Δt (Period=6) (d) Δt (Period=7)

Alg.	**Sine**	**MG**	**Lorenz**
CNN	46.0 (8.22)	12.8 (3.92)	9.90 (3.33)
CNNT	20.2 (7.65)	3.50 (1.29)	5.97 (2.41)
CNNT-th	8.44 (4.58)	3.00 (1.21)	8.37 (3.24)
ICNN-L	1.13 (0.87)	0.97 (0.53)	5.81 (2.78)
ICNN-Q	0.75 (0.65)	0.83 (0.46)	5.08 (2.59)
ICNN-C	0.72 (0.83)	0.72 (0.42)	4.22 (2.27)
ICNN-P	20.5(6.43)	1.95(0.79)	8.50(3.32)
ICNN-S	17.2(5.57)	3.51(1.36)	8.20(3.31)
RNNT	36.1(12.9)	8.15(3.32)	13.4(3.95)
RNNT-th	19.5(6.48)	8.48(3.11)	13.9(4.36)
CCNN	0.88 (0.61)	2.46 (0.89)	3.93 (1.73)
CCNN-th	**0.42** (0.36)	**0.53** (0.97)	**3.25** (1.67)

MSE loss in 10^{-2}

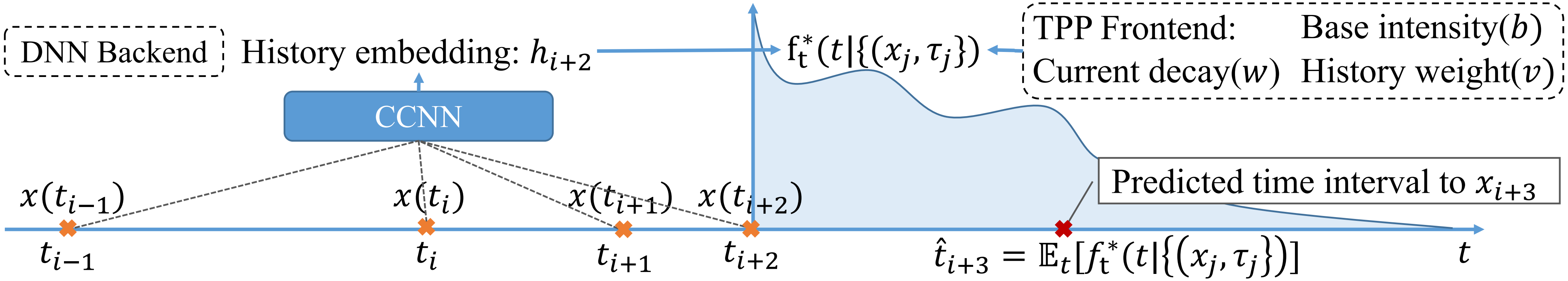
Evaluations: Temporal Point Process



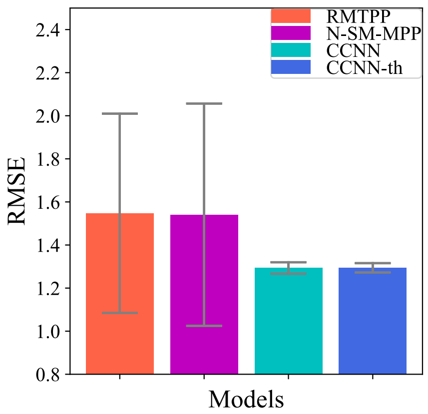
$\lambda(\tau) = \exp(vh + w\tau + b)$

$f(t) = \int^t \lambda(\tau) d\tau$

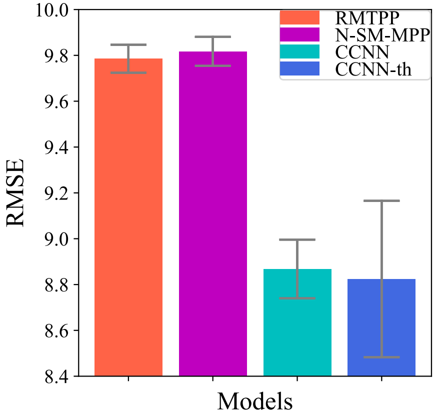
$\Delta T = \mathbb{E}(f(t))$



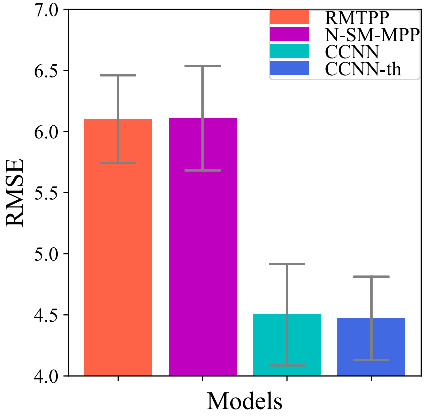
Evaluations: Temporal Point Process



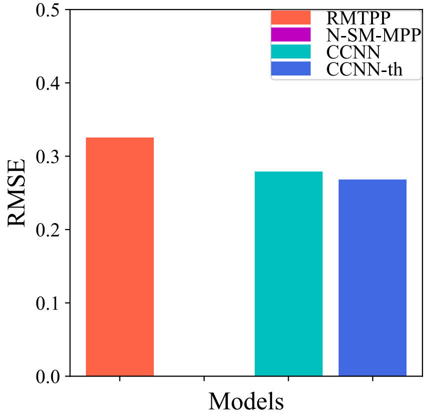
NYSE



StackOverflow



MIMIC



ReTweet

Learn more about CCNN

Paper Link

Code Link

Thank you!