SONATA: Synergistic Coreset Informed Adaptive Temporal Tensor Factorization

ICLR 2026

Maolin Wang*, Zhiqi Li, Binhao Wang, Xuhui Chen, Tianshuo Wei, Wanyu Wang, Shikai Fang, Ruocheng Guo, Zenglin Xu†, Xiangyu Zhao†

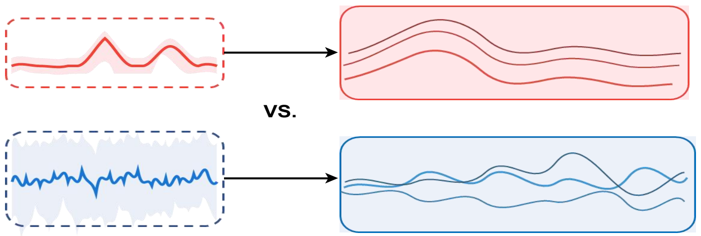
City University of Hong Kong • Zhejiang University • Fudan University • Shanghai Academy of AI for Science • ★ Code: github.com/Applied-Machine-Learning-Lab/ICLR2026_SONATA

Motivation

Dynamic tensor streams arise in recommender systems, neuroscience, and spatiotemporal sensing. Two fundamental challenges persist:

Modeling Expressiveness
Existing methods oversimplify temporal representations and fail to capture complex, evolving multi-scale entity dynamics.

Stream Efficiency
Processing all observations is computationally prohibitive. Most data is redundant; a small fraction is disproportionately informative.



Problem Formulation

Observe K-mode tensor stream {(ℓ , y , t)}. Learn dynamic embeddings $u^{(k)}_j(t) : \mathbb{R}^+ \to \mathbb{R}^R$ encoding evolving entity properties via CP factorization:

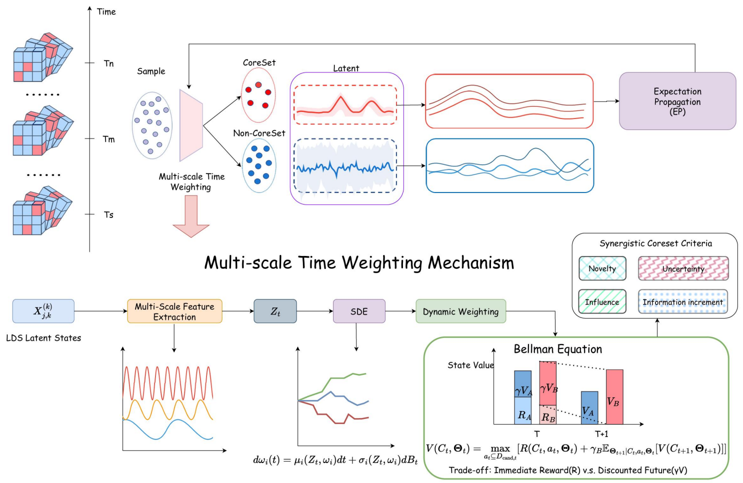
$$y_n = \sum_{r=1}^{R} \prod_{k=1}^{K} u^{(k)}_{l_{n,k},r}(t_n) + \epsilon_n := f(\{\mathbf{u}^{(k)}_{l_{n,k}}(t_n)\}_{k=1}^{K}) + \epsilon_n$$

Entity evolution via Stochastic Differential Equations (Matérn kernels ⇒ LDS):

$$d\boldsymbol{x}_j^{(k)}(t) = \mathbf{F}\boldsymbol{x}_j^{(k)}(t)dt + \mathbf{L}d\boldsymbol{w}(t) \qquad \boldsymbol{u}_j^{(k)}(t) = \mathbf{H}\boldsymbol{x}_j^{(k)}(t)$$

SONATA Framework Overview

SONATA unifies expressive LDS-based temporal modeling with principled synergistic coreset selection and streaming Bayesian inference — without deep neural networks.



Synergistic Coreset Criteria

$$S_n = w_u \cdot \mathcal{I}_{\text{unc}}(n) + w_i \cdot \mathcal{I}_{\text{inf}}(n) + w_n \cdot \mathcal{I}_{\text{nov}}(n) + w_m \cdot \mathcal{I}_{\text{mart}}(n)$$

Uncertainty I_{unc}	Influence I_{inf}
Novelty I_{nov}	Info Increment I_{mart}

Bellman-Inspired Coreset Evolution

$$V(\mathcal{C}_t, \boldsymbol{\Theta}_t) = \max_{a_t \subseteq \mathcal{D}_{\text{cand},t}} \left[\mathcal{R}(\mathcal{C}_t, a_t, \boldsymbol{\Theta}_t) + \gamma_B \mathbb{E}_{\boldsymbol{\Theta}_{t+1}|\mathcal{C}_t, a_t, \boldsymbol{\Theta}_t}[V(\mathcal{C}_{t+1}, \boldsymbol{\Theta}_{t+1})]\right]$$

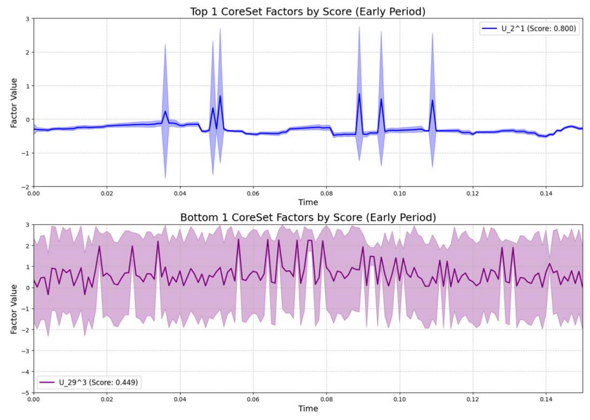
Balances immediate reward (importance score) vs. discounted future utility. Enables strategic data retention outperforming myopic selection.

Experimental Results (RMSE, R=5, 10 runs)

Method	CA Traffic	ServerRoom	BeijingAir	FitRecord
NONFAT	0.501	0.117	0.395	0.503
THIS-ODE	0.632	0.132	0.540	0.526
SFTL-CP	0.231	0.161	0.248	0.424
SONATA (Ours)	**0.089***	**0.115***	**0.237***	**0.414***

61.5% RMSE reduction on CA Traffic 30K

Visualization of SONATA's CoreSet selection effectiveness



Conclusion

- Novel framework: LDS temporal modeling + synergistic coreset selection
- 4-criterion scoring: uncertainty, influence, novelty, info increment
- Bellman-guided long-term optimization beyond greedy selection
- 25× speedup over full-data, SOTA on 4 streaming tensor datasets