Proceedings of 41th International Conference on Machine Learning (ICML 2024)

TimeSiam: A Pre-Training Framework for Siamese Time-Series Modeling

Jiaxiang Dong[*,1] Haixu Wu[*,1] Yuxuan Wang[1] Yunzhong Qiu[1] Li Zhang[1] Jianmin Wang[1] Mingsheng Long[1]

[*]Equal contribution [1]School of Software, BNRist, Tsinghua University.
Jiaxiang Dong <djx20@mails.tsinghua.edu.cn>. Haixu Wu <wuhx23@mails.tsinghua.edu.cn>
Correspondence to: Mingsheng Long <mingsheng@tsinghua.edu.cn>


Jiaxiang Dong


Haixu Wu


Yuxuan Wang


Yunzhong Qiu


Li Zhang

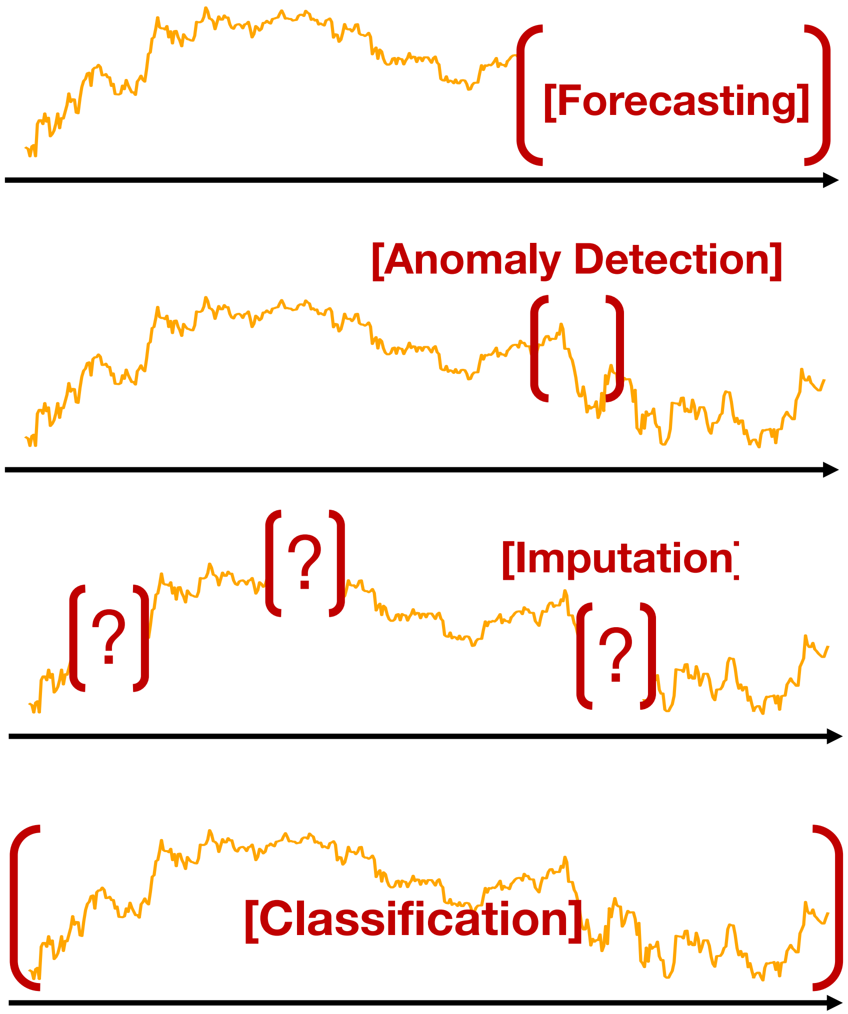

Jianmin Wang


Mingsheng Long

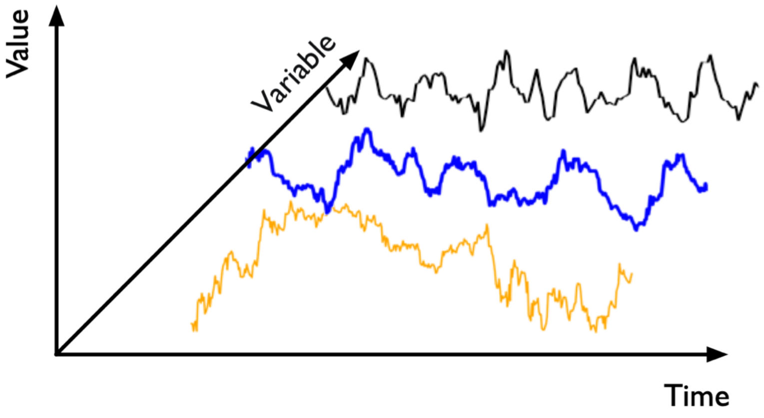
Time Series In Real World

Data

Task

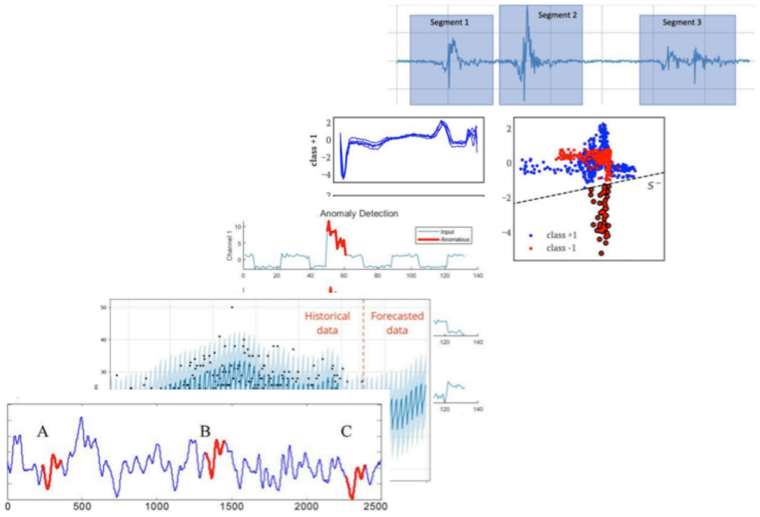


Pre-training and Fine-tuning in Time Series



Pre-training

Fine-tuning



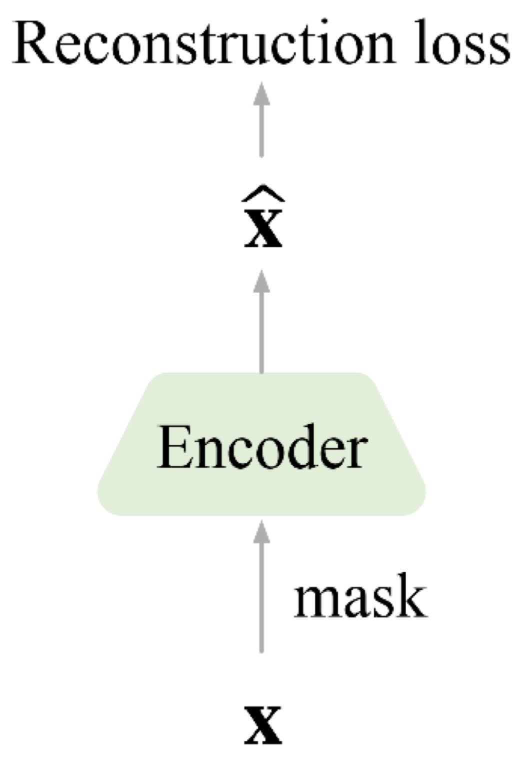
Diverse time series data

Downstream time series analysis tasks

① Use the model as the carrier of knowledge.

② Learn transferable temporal representations.

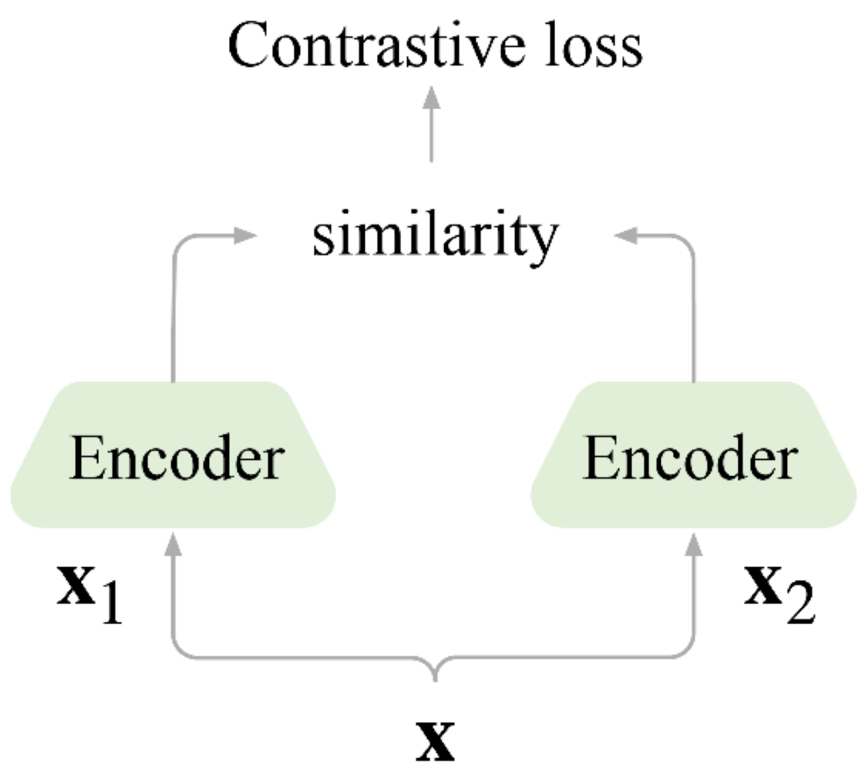
Pre-training Methods in Time Series



Masked Modeling

Reconstruct the masked content based on the unmasked part.

（TST、PatchTST、SimMTM…）

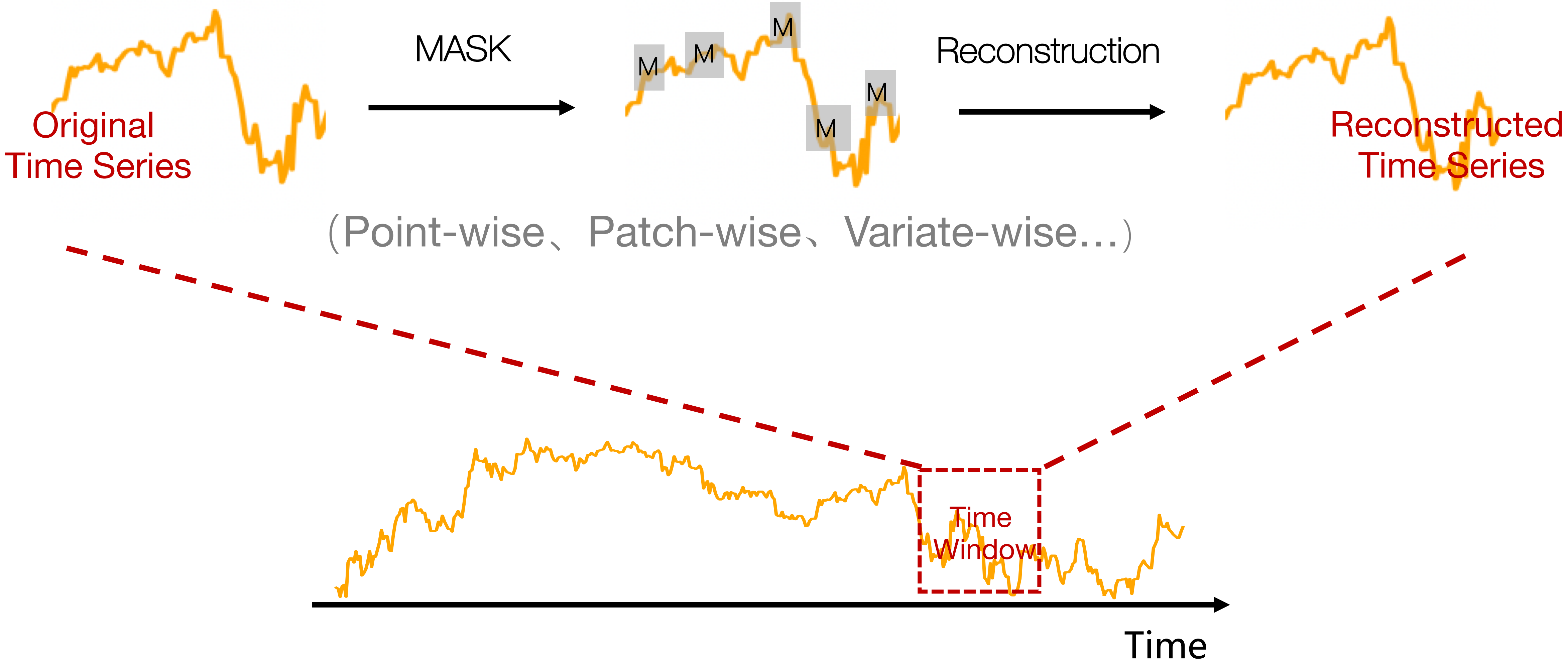


Contrastive Learning

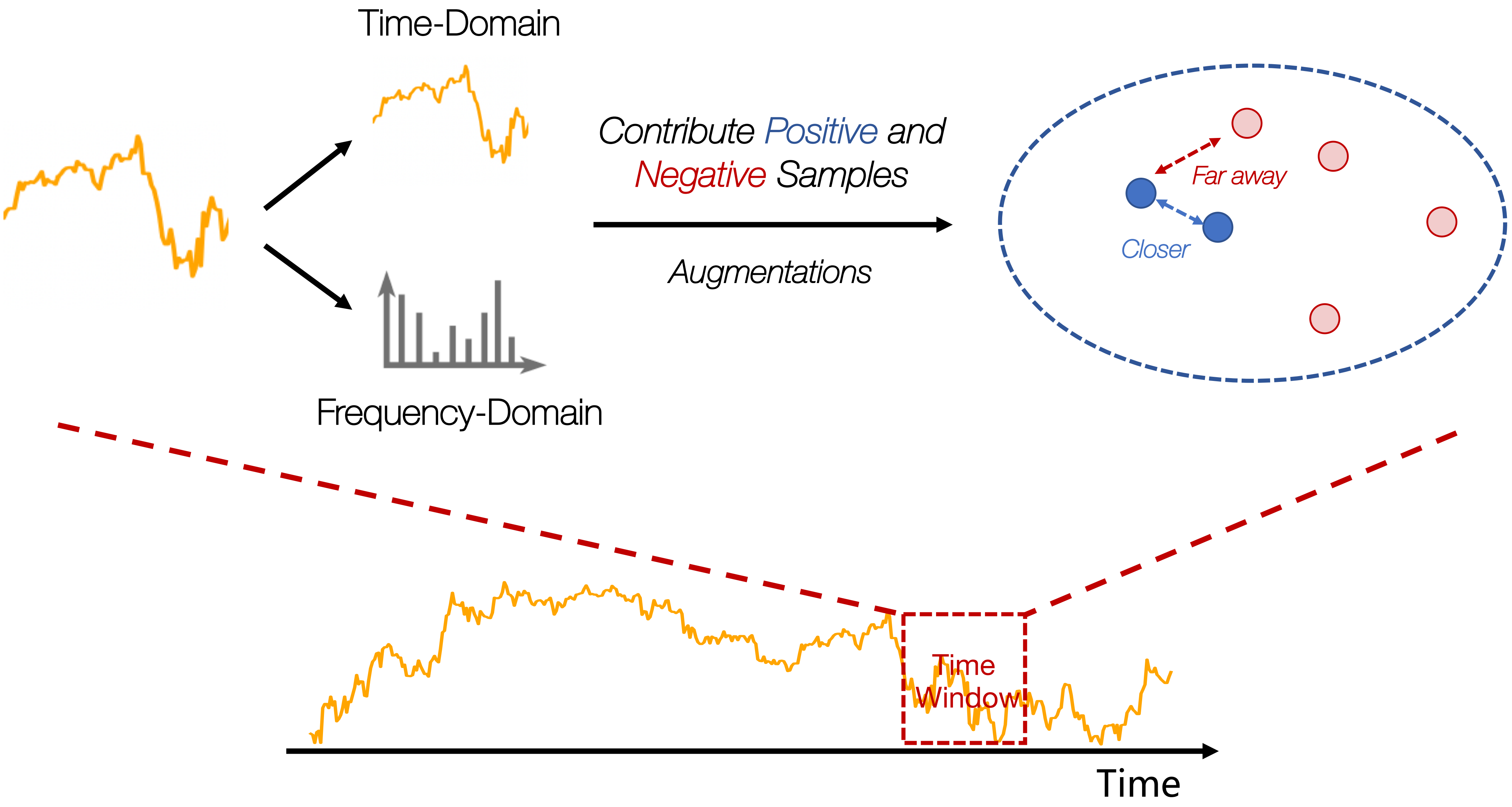
Learn discriminative positive or negative representations.

（TS2Vec、TFC、COMET…）

Masked Modeling in Time Series



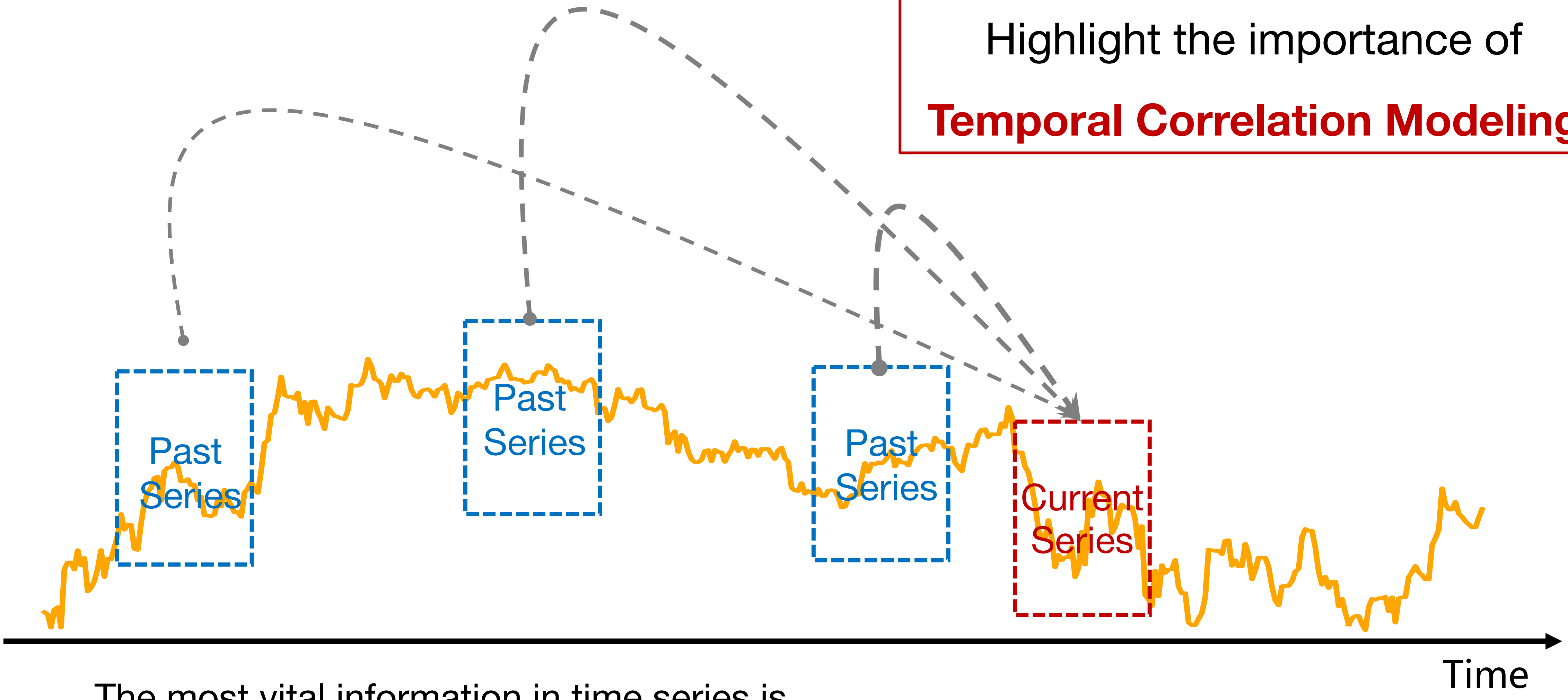
Contrastive Learning in Time Series



Temporal Correlation Modeling in Time Series

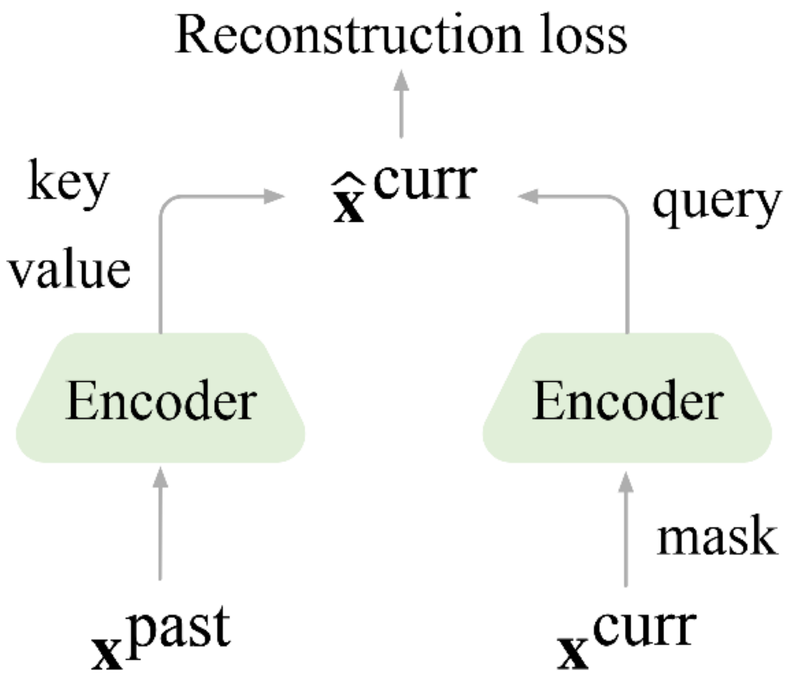
Highlight the importance of

Temporal Correlation Modeling



The most vital information in time series is preserved in the temporal correlations.

Siamese Masked Modeling



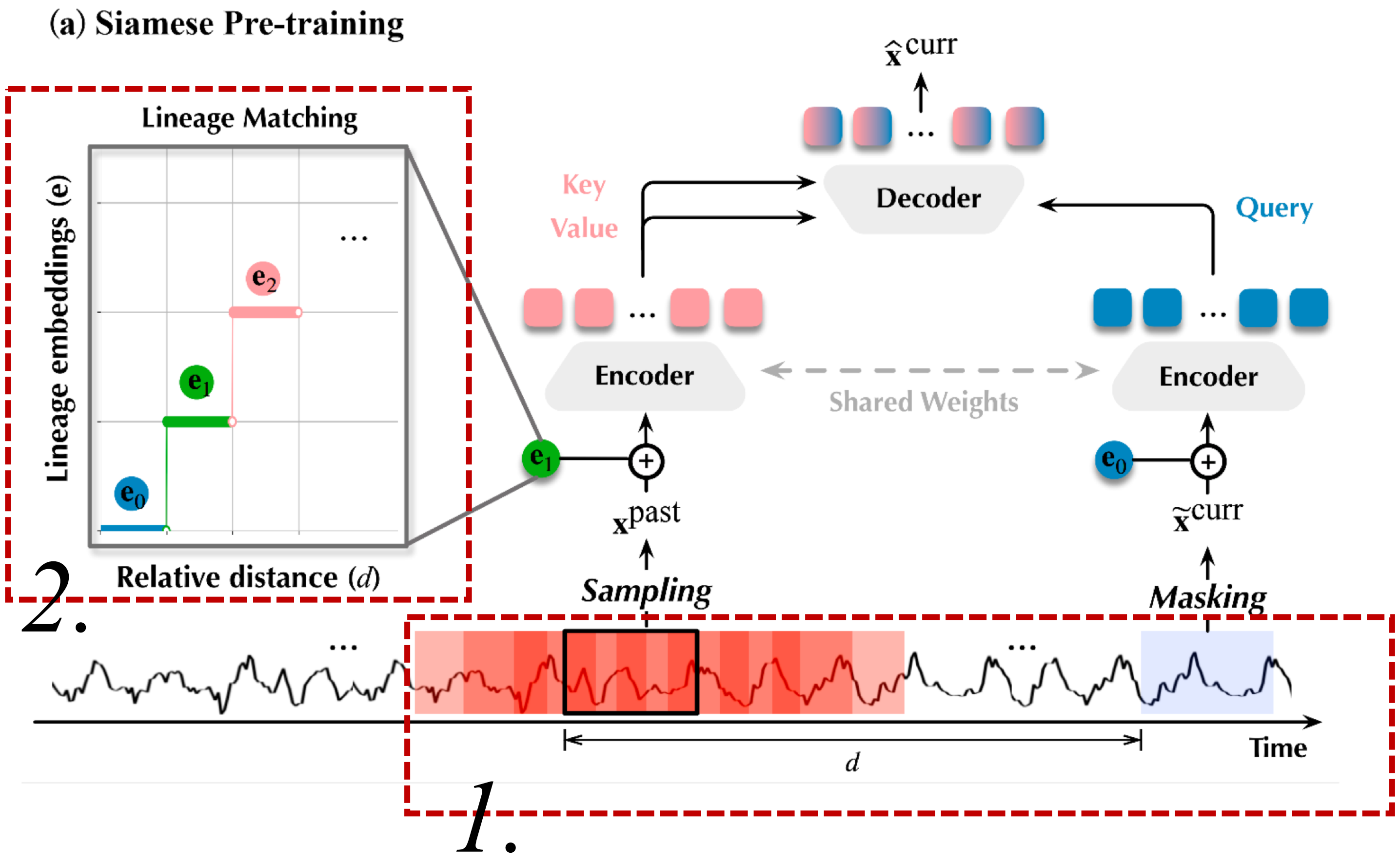
Emphasis on modeling time series association relationships from the past to the current

① Siamese Network & Subseries

② Past-to-Current Reconstruction

③ Learnable Lineage Embeddings

Overall design of TimeSiam



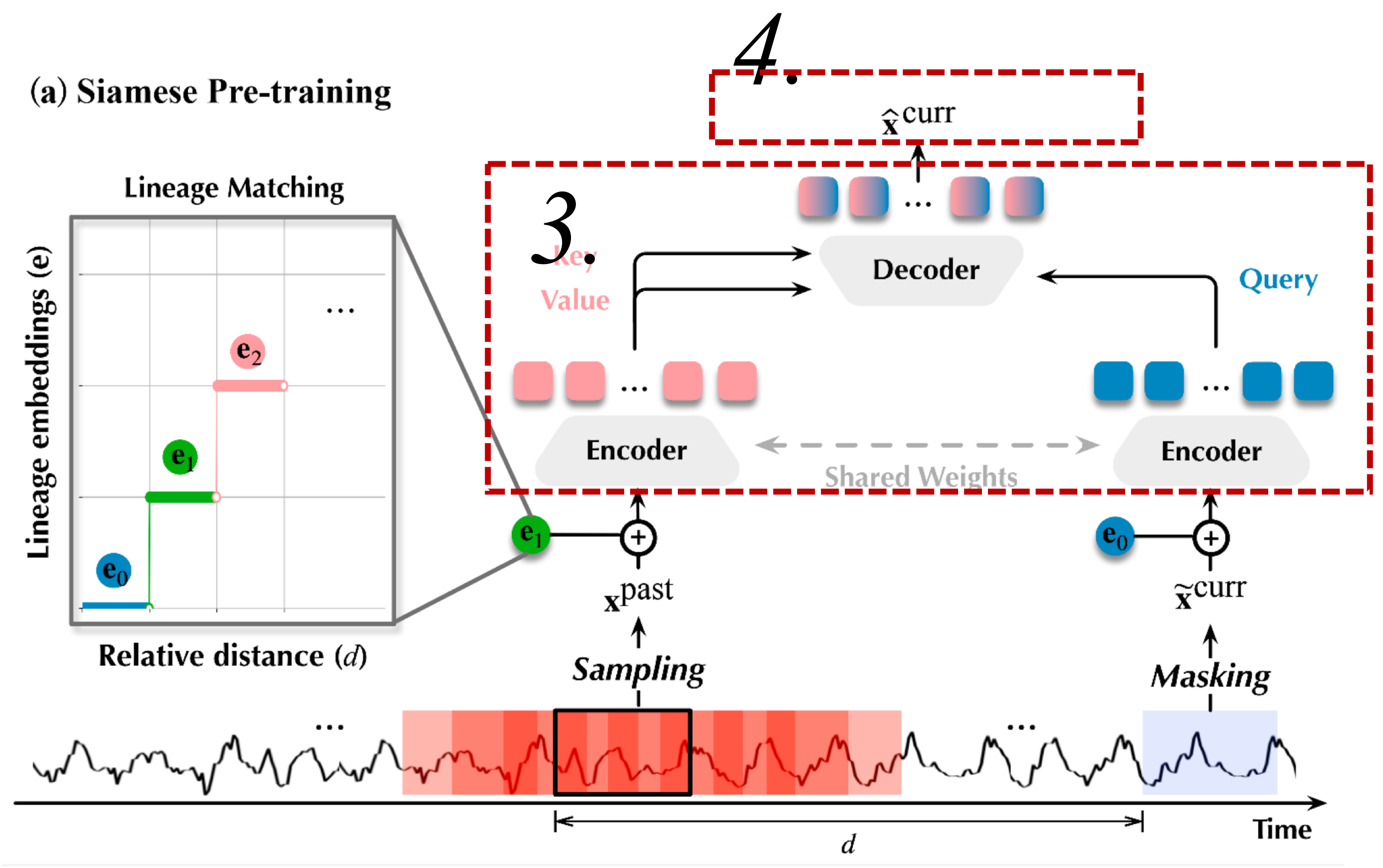
① Siamese Sampling

$$(\mathbf{x}^{\text{past}}, \widetilde{\mathbf{x}}^{\text{curr}}) = \text{Mask-Augment}\left((\mathbf{x}^{\text{past}}, \mathbf{x}^{\text{curr}})\right). \quad (1)$$

② Lineage Matching

$$\begin{aligned} \mathbf{e}_i^{\text{lineage}} &= \text{LineageMatching}(d) \\ \mathbf{z}^{\text{past}} &= \text{Embed}(\mathbf{x}^{\text{past}}) \oplus \mathbf{e}_i^{\text{lineage}} \\ \widetilde{\mathbf{z}}^{\text{curr}} &= \text{Embed}(\widetilde{\mathbf{x}}^{\text{curr}}) \oplus \mathbf{e}_0^{\text{lineage}}, \end{aligned} \quad (2)$$

Overall design of TimeSiam



③ Representation Learning & Aggregation

$$\mathbf{h}_e^{\text{past}} = \text{Encoder}(\mathbf{z}^{\text{past}}),\ \tilde{\mathbf{h}}_e^{\text{curr}} = \text{Encoder}(\tilde{\mathbf{z}}^{\text{curr}}), \tag{3}$$

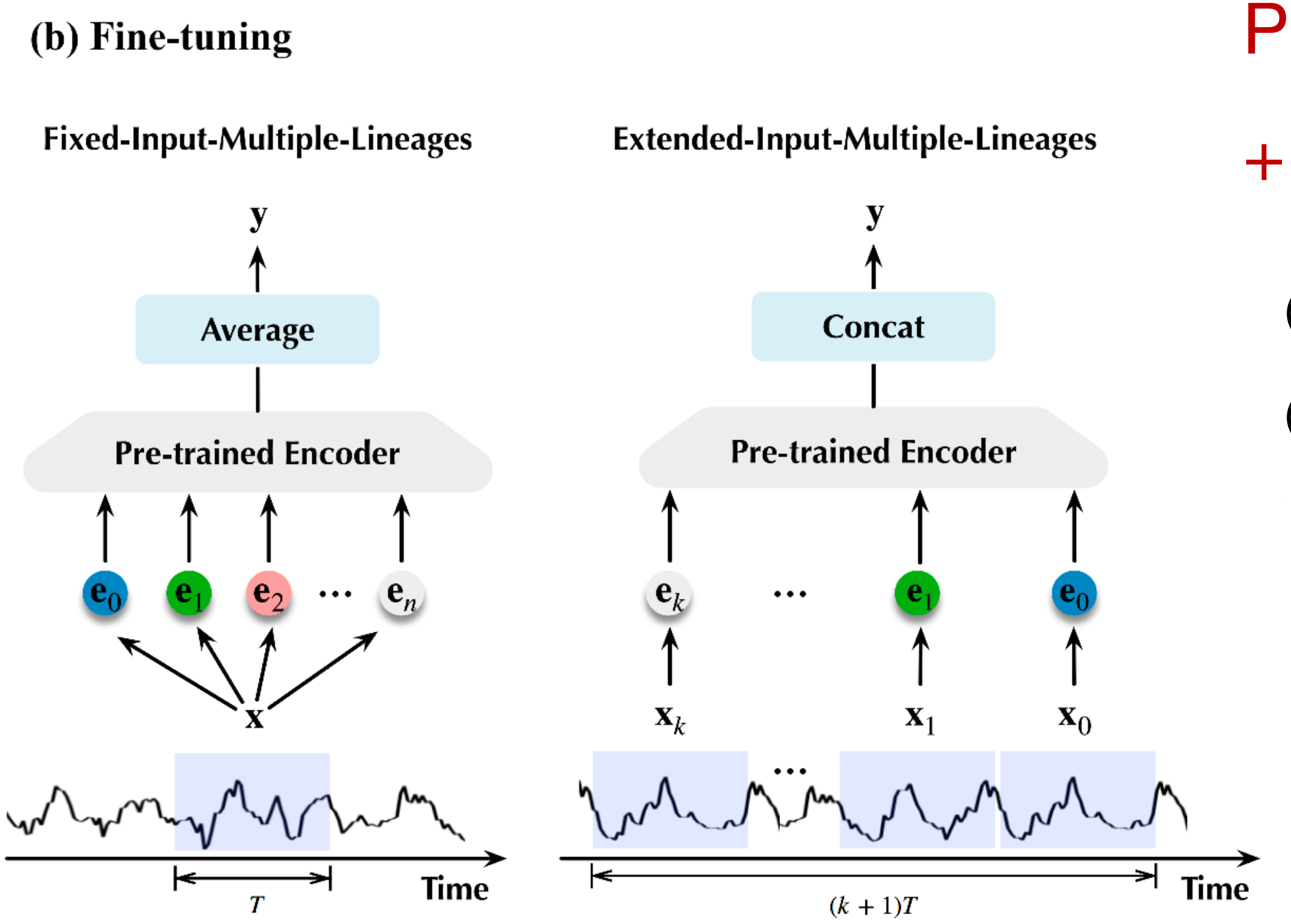
$$\begin{aligned}
\hat{\mathbf{h}}_d &= \text{LayerNorm}\left(\tilde{\mathbf{h}}_e^{\text{curr}} + \text{Cross-Attn}\left(\tilde{\mathbf{h}}_e^{\text{curr}}, \mathbf{h}_e^{\text{past}}, \mathbf{h}_e^{\text{past}}\right)\right) \\
\mathbf{h}'_d &= \text{LayerNorm}\left(\hat{\mathbf{h}}_d + \text{Self-Attn}\left(\hat{\mathbf{h}}_d, \hat{\mathbf{h}}_d, \hat{\mathbf{h}}_d\right)\right) \\
\mathbf{h}_d &= \text{LayerNorm}\left(\mathbf{h}'_d + \text{FFN}\left(\mathbf{h}'_d\right)\right).
\end{aligned} \tag{4}$$

④ Series Reconstruction

$$\hat{\mathbf{x}}^{\text{curr}} = \text{Projector}(\mathbf{h}_d). \tag{5}$$

$$\mathcal{L}_{\text{reconstruction}} = \|\mathbf{x}^{\text{curr}} - \hat{\mathbf{x}}^{\text{curr}}\|_2^2. \tag{6}$$

Overall design of TimeSiam



Pretrained Siamese Networks

+ Diverse Lineage Embeddings

1. Enrich Fixed representation
2. Enhance Extended Representations

$$\begin{aligned}\overline{\mathbf{h}}_e &= \mathrm{Average}\left(\mathbf{h}_{e,0}, \mathbf{h}_{e,1}, \ldots \mathbf{h}_{e,n}\right), \\ \text{where } \mathbf{h}_{e,i} &= \mathrm{Encoder}\left(\mathrm{Embed}(\mathbf{x}) \oplus \mathbf{e}_i^{\text{lineage}}\right).\end{aligned} \tag{7}$$

$$\begin{aligned}\overline{\mathbf{h}}_e &= \mathrm{Concat}\left(\mathbf{h}_{e,0}, \mathbf{h}_{e,1}, \ldots, \mathbf{h}_{e,k}\right), \\ \text{where } \mathbf{h}_{e,i} &= \mathrm{Encoder}\left(\mathrm{Embed}(\mathbf{x}_i) \oplus \mathbf{e}_{\text{LineageMatching}(iT)}^{\text{lineage}}\right).\end{aligned} \tag{8}$$

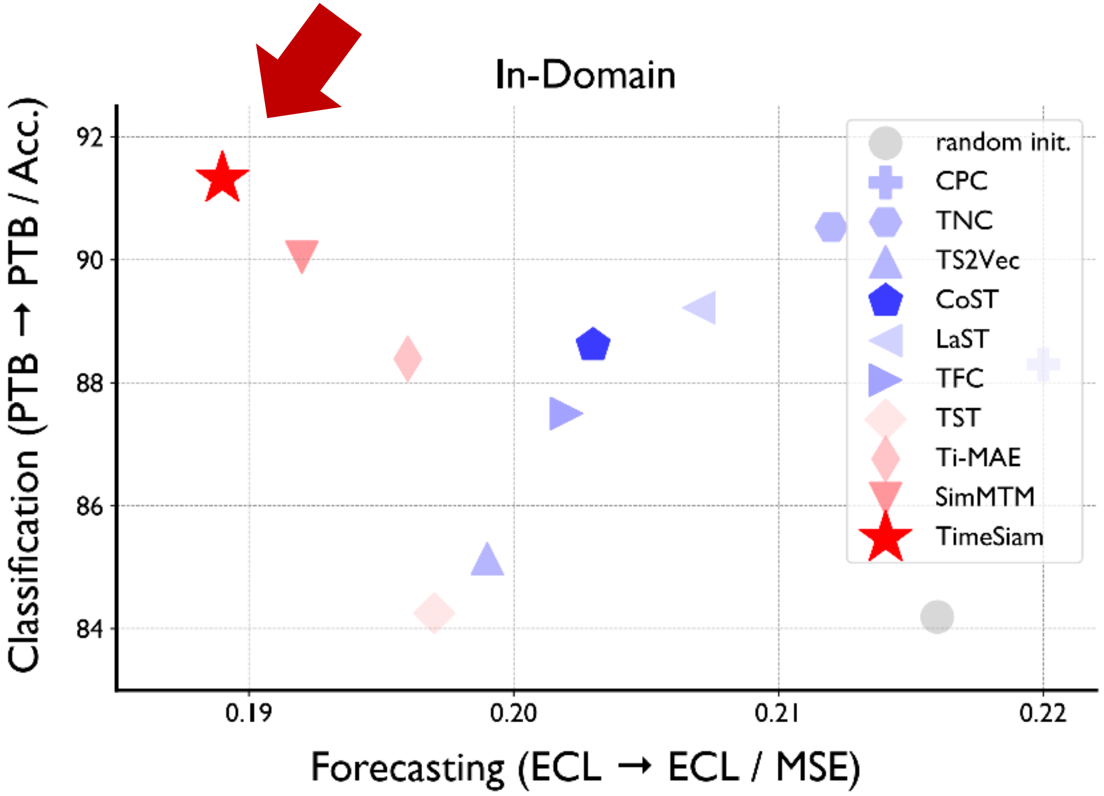
Experiment: Overall

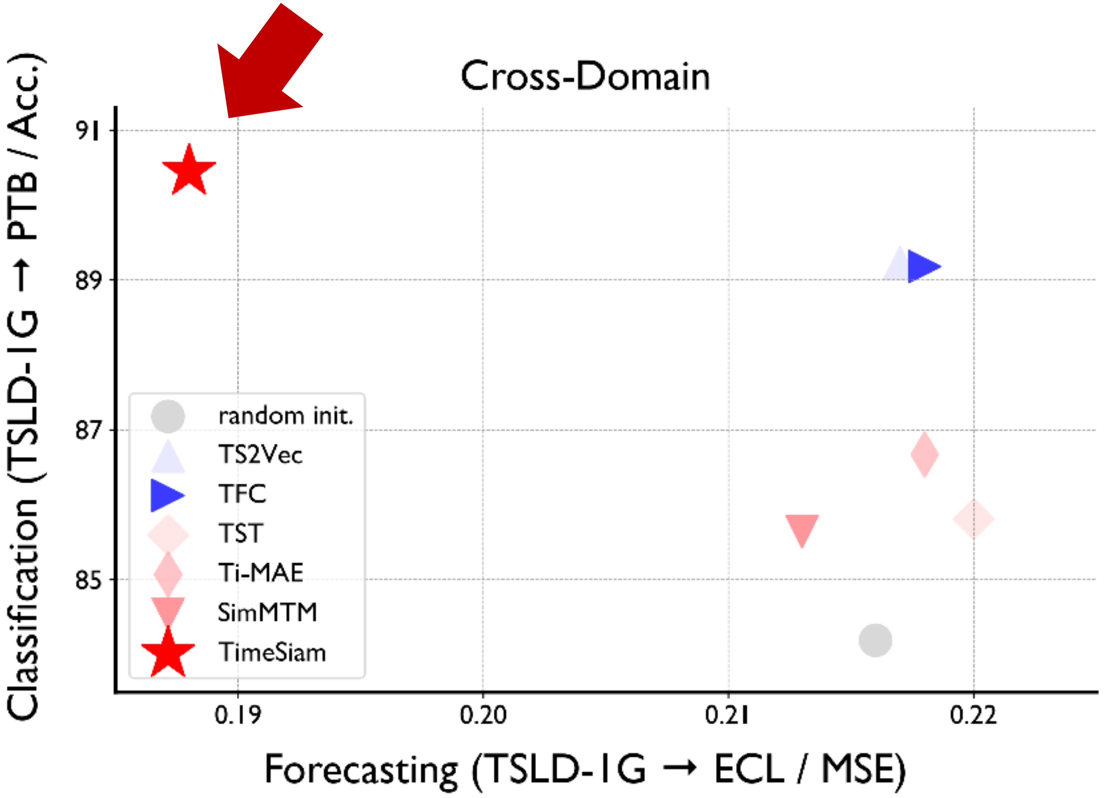
Table 1. Summary of experiment benchmarks, where TSLD-500M and TSLD-1G are newly constructed from diverse domains.

TASKS	DATASETS	DOMAIN	EXAMPLES
Forecasting	ETT (4 subsets)	Electricity	14.3K
	Weather	Weather	52.7K
	Electricity	Electricity	26.3K
	Traffic	Transportation	17.5K
	Exchange	Finance	7.6K
	TSLD-500M	Multiple	412.6K
	TSLD-1G	Multiple	13.9M
Classification	AD	EEG	5.97K
	TDBrain	EEG	11.9K
	PTB	ECG	62.4K

- ✓ Two typical time series analysis tasks: **Forecasting and Classification.**
- ✓ Under multiple experiment settings: **In- and Cross domain， Single and Multiple Mixed Datasets.**
- ✓ Compared to **11 advanced baselines covering 13 standard benchmarks.**

Experiment: Overall





TimeSiam outperforms other baselines significantly in all settings!

Experiment: Forecasting and Classification

Encoder	Method	ETTh1	ETTh2	ETTm1	ETTm2	Weather	Exchange	ECL	Traffic
PatchTST	Random Init.	0.473	0.385	0.390	0.285	0.259	0.367	0.216	0.490
	CPC (2018)	0.440	0.401	0.389	0.290	0.272	0.368	0.220	0.504
	TNC (2021)	0.445	0.379	0.386	0.287	0.270	0.362	0.212	0.501
	TS2Vec (2022)	0.456	0.376	0.393	0.289	0.256	0.363	0.199	0.472
	CoST (2022)	0.457	0.374	0.395	0.286	0.253	0.364	0.203	0.480
	LaST (2022)	0.479	0.385	0.398	0.285	0.252	0.433	0.207	0.520
	TF-C (2022)	0.453	0.378	0.389	0.281	0.257	0.362	0.202	0.487
	TST (2021)	0.452	0.383	0.380	0.288	0.259	0.385	0.197	0.486
	Ti-MAE (2023b)	0.448	0.379	0.384	0.279	0.257	0.370	0.196	0.481
	PatchTST† (2023)	0.442	0.381	0.379	0.285	0.267	0.358	0.200	0.484
	SimMTM (2023)	0.440	0.382	0.377	0.285	0.256	0.361	0.192	0.466
	TimeSiam	**0.429**	**0.373**	**0.374**	**0.279**	**0.252**	**0.353**	**0.189**	**0.453**
iTransformer	Random Init.	0.454	0.383	0.407	0.288	0.258	0.365	0.178	0.428
	TS2Vec (2022)	0.474	0.379	0.411	0.290	0.264	0.364	0.246	0.485
	CoST (2022)	0.472	0.386	0.411	0.294	0.269	0.366	0.252	0.529
	LaST (2022)	0.465	0.386	0.400	0.302	0.262	0.386	0.237	0.477
	TF-C (2022)	0.450	0.379	0.403	0.292	0.265	0.372	0.222	0.432
	TST (2021)	0.447	0.376	0.399	0.291	0.261	0.363	0.228	0.438
	Ti-MAE (2023b)	0.448	0.378	0.399	0.289	0.257	0.366	0.217	0.430
	SimMTM (2023)	0.445	0.376	0.397	0.286	0.259	0.358	0.179	0.426
	TimeSiam	**0.440**	**0.371**	**0.390**	**0.284**	**0.256**	**0.355**	**0.175**	**0.420**

Method	ECL	Traffic
	.216	
	.217	0.528
	.218	0.543
	.220	0.514
Ti-MAE (2023b)	0.218	0.515

Method	ETTh1	ETTh2	ETTm1	ETTm2	Weather	Exchange	ECL	Traffic
SimMTM (2023)	0.429	0.380	0.375	0.287	0.252	0.365	0.213	0.459
TimeSiam	**0.425**	**0.374**	**0.371**	0.286	**0.251**	**0.360**	**0.188**	**0.454**

Method	AD	TDBrain	PTB
Random Init.	80.62	79.08	84.19
CPC (2018)	77.40	85.19	88.30
TNC (2021)	78.58	85.21	90.53
TS2Vec (2022)	81.26	80.21	85.14
CoST (2022)	73.87	83.86	88.61
LaST (2022)	72.63	85.13	89.22
TF-C (2022)	75.31	66.62	87.50
COMET (2023)	84.50	85.47	87.84
TST (2021)	81.50	83.22	84.25
Ti-MAE (2023b)	80.70	88.16	88.39
SimMTM (2023)	86.19	84.81	90.04
TimeSiam	**89.93**	**90.67**	**91.32**

Method			PTB
			84.19
			89.23
			89.18
TST (2021)	82.60	83.65	85.81
Ti-MAE (2023b)	80.40	85.22	86.67
SimMTM (2023)	87.74	85.29	85.64
TimeSiam	**90.47**	**86.26**	**90.45**

TimeSiam consistently outperforms other pre-training methods for forecasting and classification tasks.

Experiment: Larger Data and Model Capability

Table 7. Fine-tuning performance of TimeSiam under different pre-training data and model sizes. Relative improvement over random initialization (%) is marked in green. See Appendix A.3 for details.

PRE-TRAIN	TRAFFIC	ECL
Random initiation	0.490	0.216
TimeSiam in-domain	0.453	0.189
$\text{TimeSiam}_{\text{Base}}$ $\text{TSLD}_{\text{500M}}$	0.462 (+5.7)	0.189 (+12.5)
$\text{TimeSiam}_{\text{Base}}$ TSLD_{1G}	0.454 (+7.4)	0.188 (+12.5)
$\text{TimeSiam}_{\text{Large}}$ TSLD_{1G}	0.433 (+11.6)	0.185 (+14.4)

Highlights the efficacy of TimeSiam and the positive correlation between data-model size and the final performance.

Experiment: Linear Probing

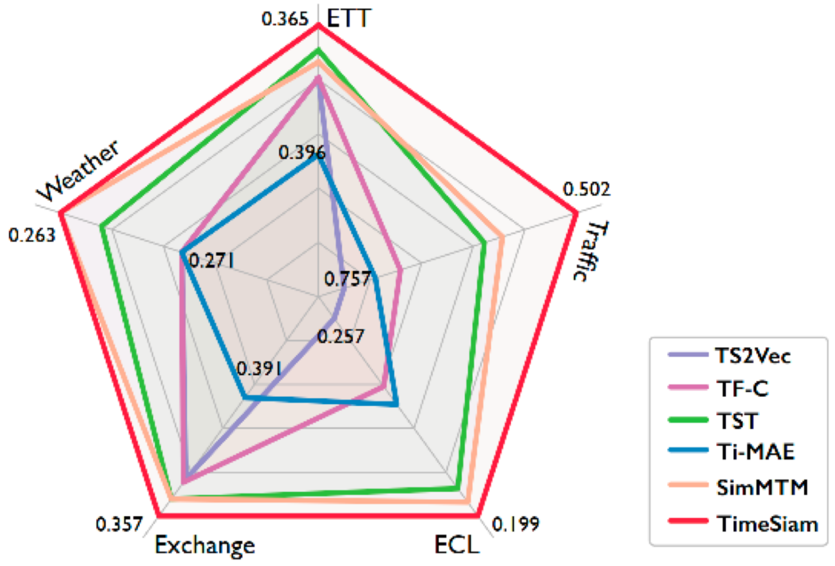


Figure 5. Linear probing on in-domain forecasting setting. Average results (MSE) are reported. Full results are shown in Table 17.

- **Higher quality temporal representations.**
- **The effectiveness in capturing the essential characteristics of time series data**

Experiment: Extended Input Length for Fine-tuning

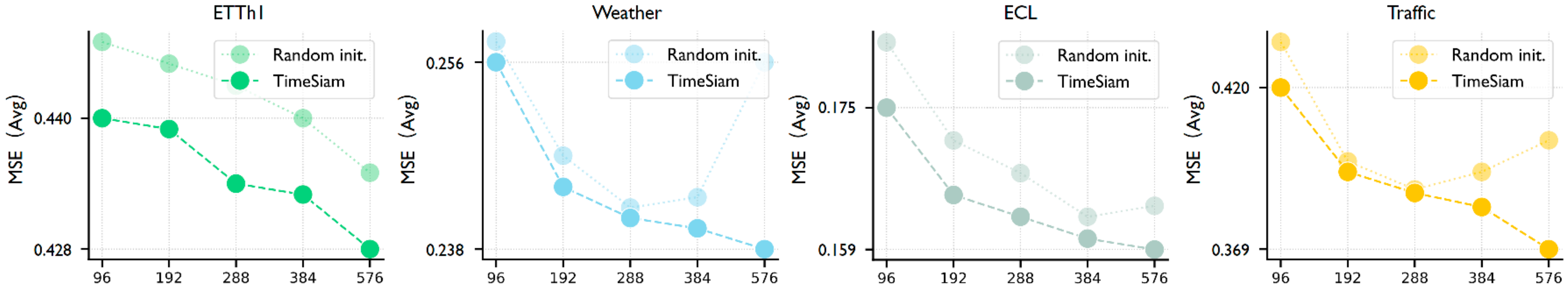
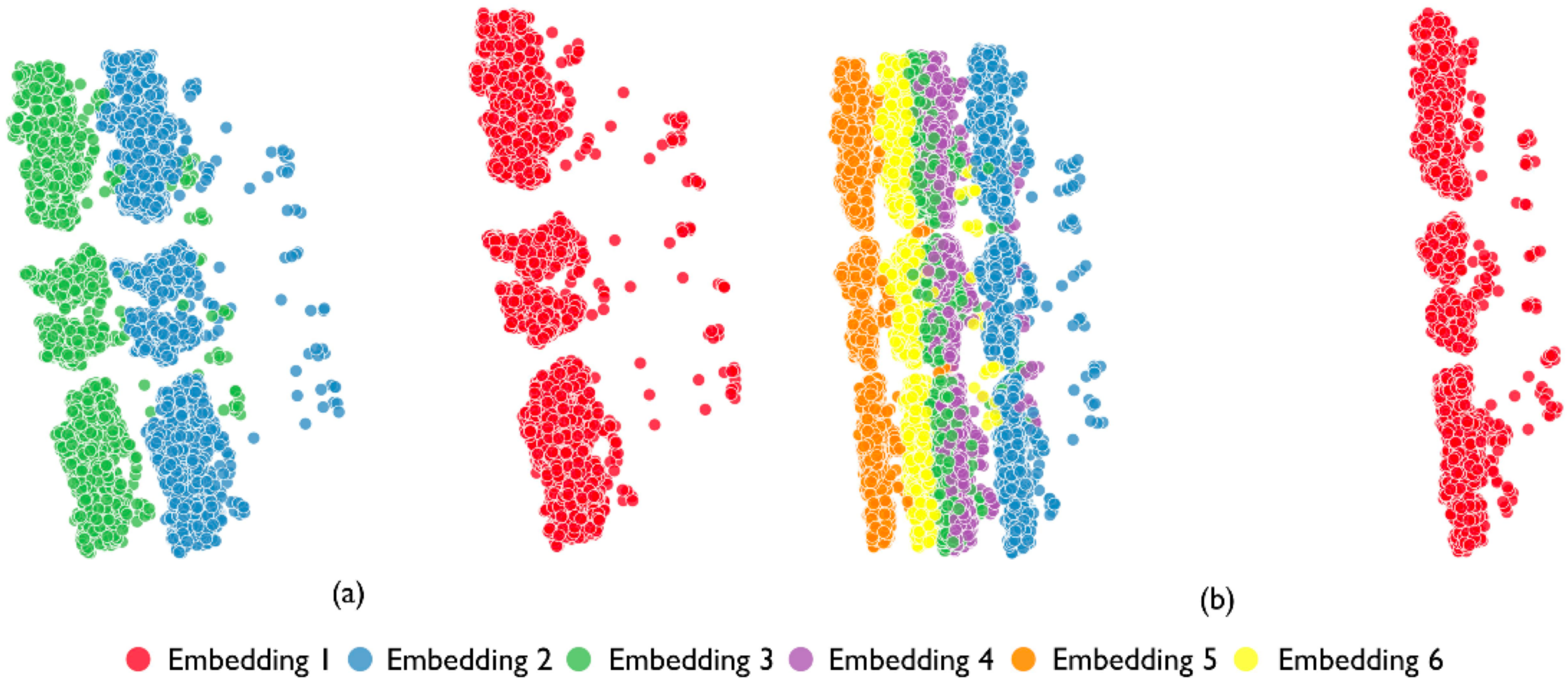


Figure 4. Fine-tuning the pre-trained model to the inputs with extended length {96, 192, 288, 384, 576} based on iTransformer (Liu et al., 2024). The MSE averaged from all predicted horizons {96, 192, 336, 720} is reported. Additional results are in the Appendix F.

Benefiting from an ingenious integration of Siamese modeling and lineage embeddings, TimeSiam achieves more accurate predictions from extended input series.

Experiment: Representation Shows



The representations generated based on the same data but with different lineage embeddings exhibit a high level of diversity

Experiment: Past-to-Current Reconstruction Shows

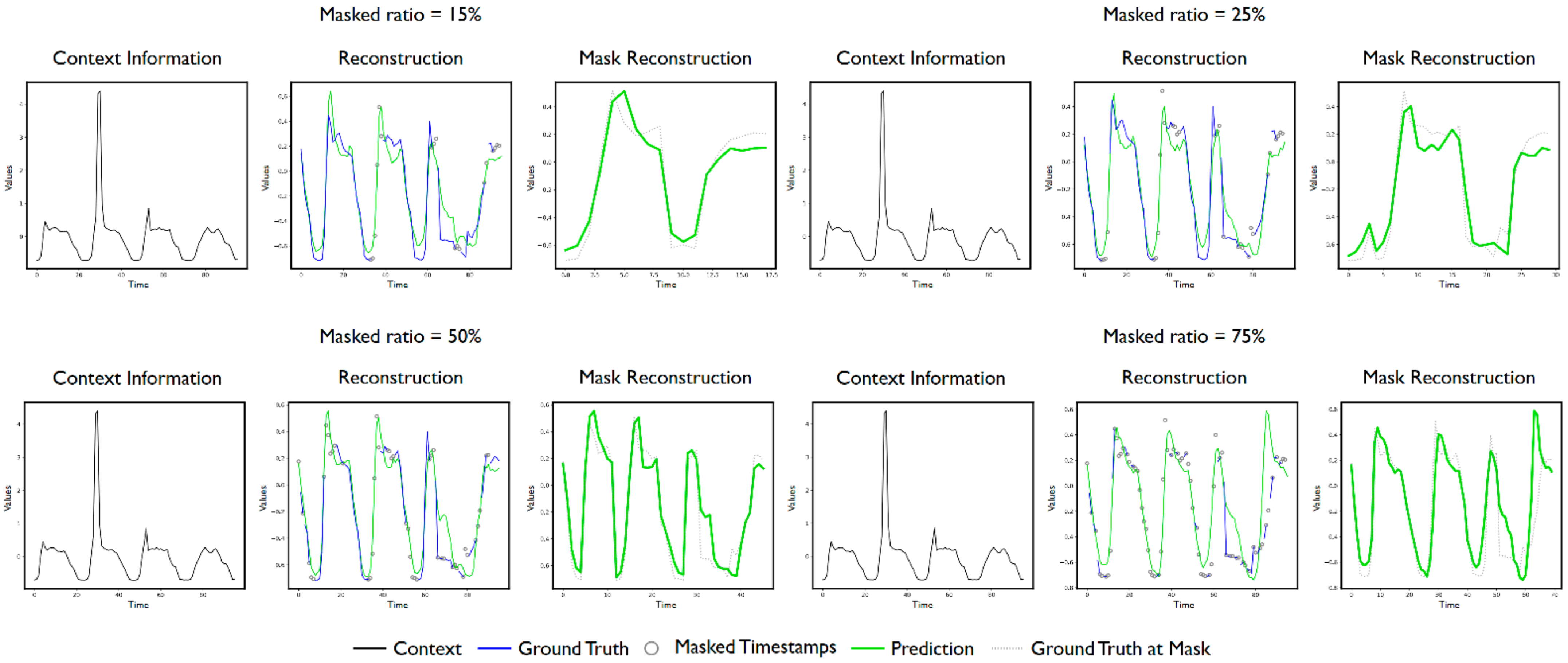
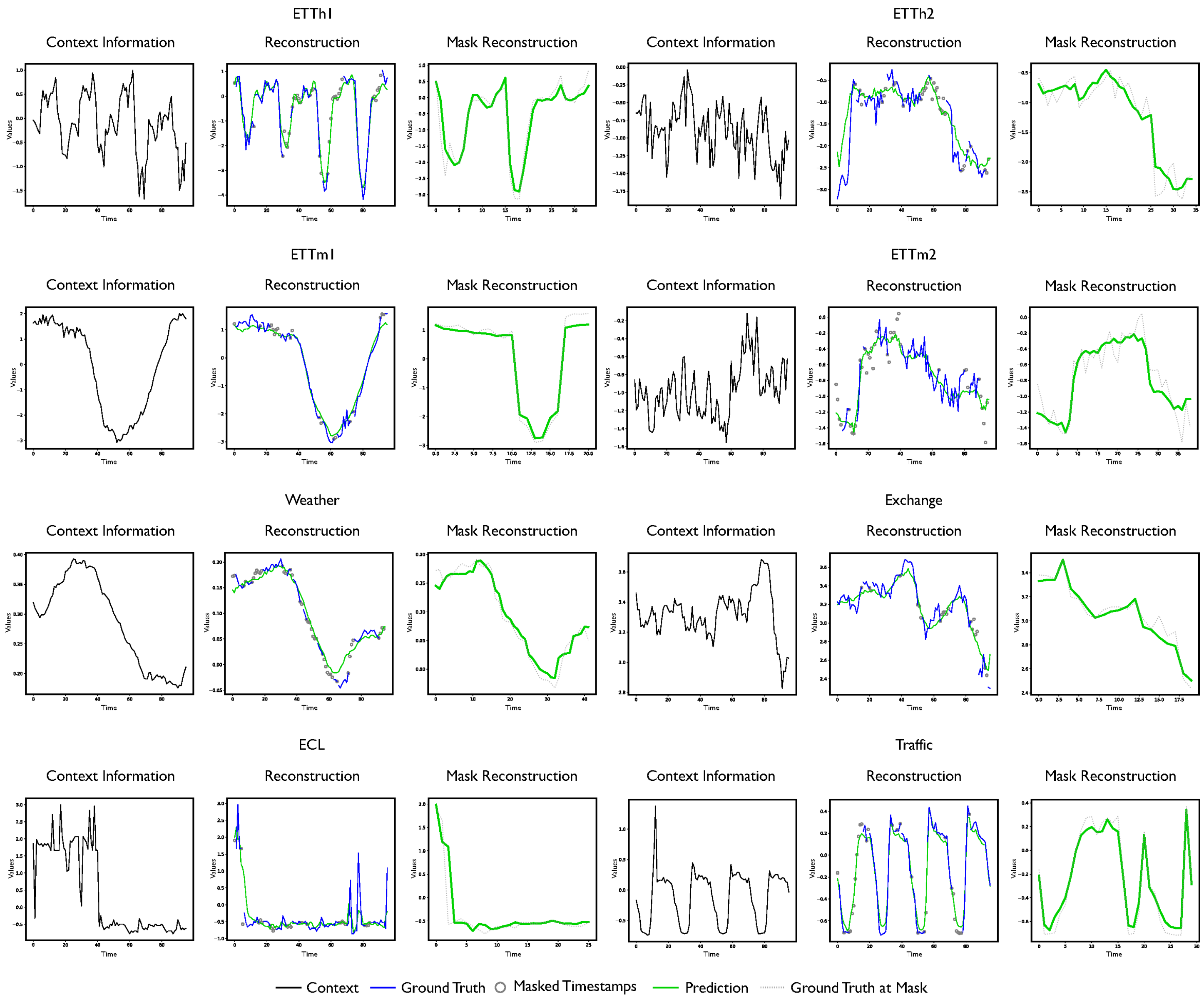


Figure 11. Showcases of TimeSiam in reconstructing time series with different masked ratios from Traffic.

Experiment: Past-to-Current Reconstruction Shows



Open Source

github.com/thuml/TimeSiam

系统网址 github 收藏网址 清华相关 工具 个人

README

0 stars
3 watching
0 forks

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

TimeSiam (ICML 2024)

This is the codebase for the paper: [TimeSiam: A Pre-Training Framework for Siamese Time-Series Modeling]

Architecture

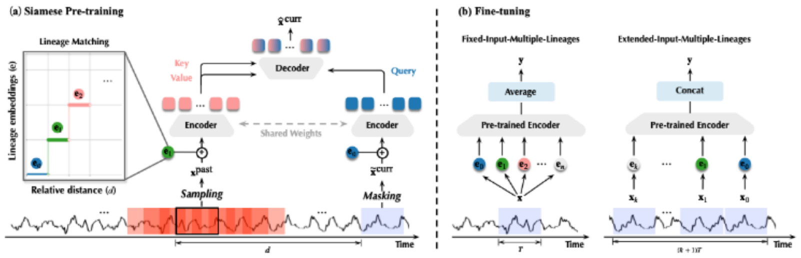


Figure 1. Overview of TimeSiam.

Citation

If you find this repo useful, please cite our paper.

```
@article{dong2024timesiam,
  title={TimeSiam: A Pre-Training Framework for Siamese Time-Series Modeling},
  author={Dong, Jiaxiang and Wu, Haixu and Wang, Yuxuan and Qiu, Yunzhong and Zhang, Li and Wa
  journal={arXiv preprint arXiv:2402.02475},
  year={2024}
}
```

Contact

We plan to open source the code in the future. If you have any questions or want to use the code, please contact djx20@mails.tsinghua.edu.cn.

Code will be open source at https://github.com/thuml/TimeSiam

Thank You!

djx20@mails.tsinghua.edu.cn

长按关注，获取最新资讯