



2023 CCF国际AIOps挑战赛决赛
暨“大模型时代的AIOps”研讨会

Foundation Models for Time Series Analysis

Haixu Wu, School of Software, Tsinghua University

主办单位：中国计算机学会（CCF）、清华大学、中国建设银行股份有限公司、南开大学

承办单位：中国计算机学会互联网专委会、清华大学计算机科学与技术系、中国建设银行股份有限公司运营数据中心、南开大学软件学院、北京必示科技有限公司

赞助单位：华为技术有限公司、国网宁夏电力有限公司电力科学研究院、软通动力信息技术（集团）股份有限公司

Time Series In Real World



Energy Consumption



Traffic Flow



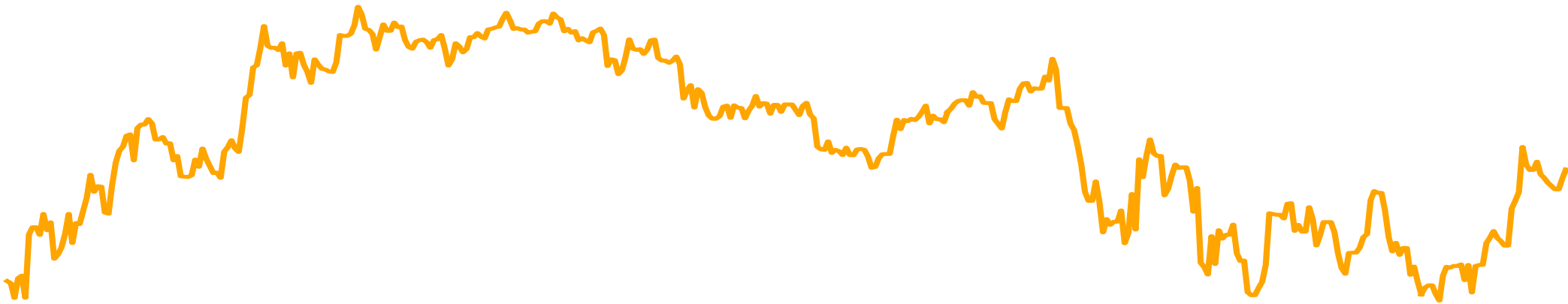
Economic Changes



Weather Variations

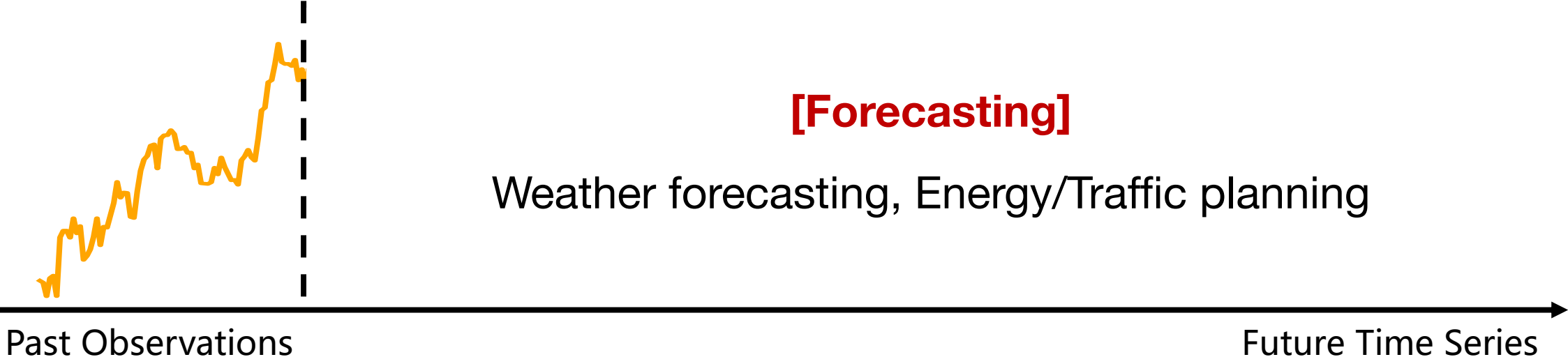


Disease Propagation

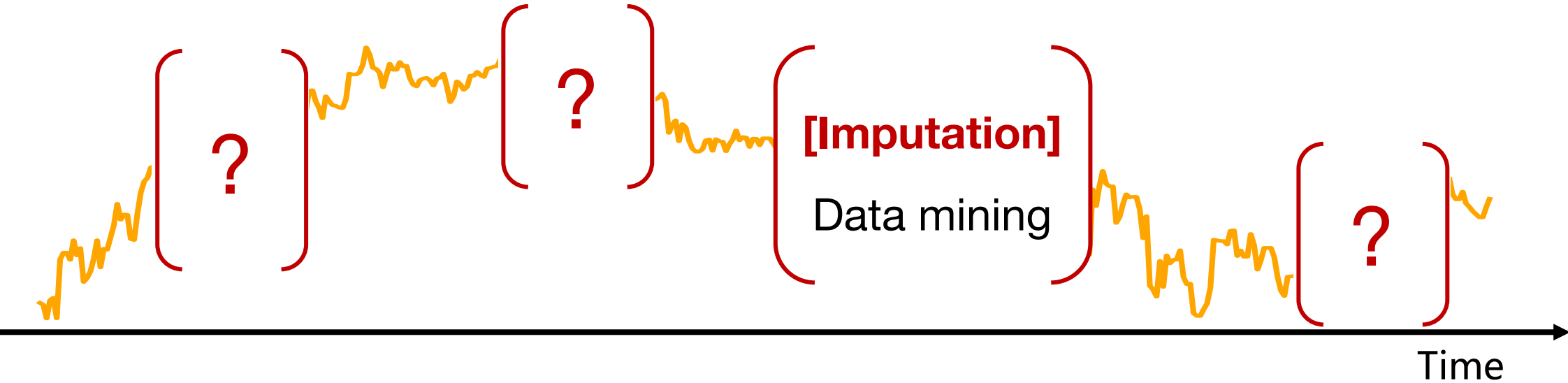
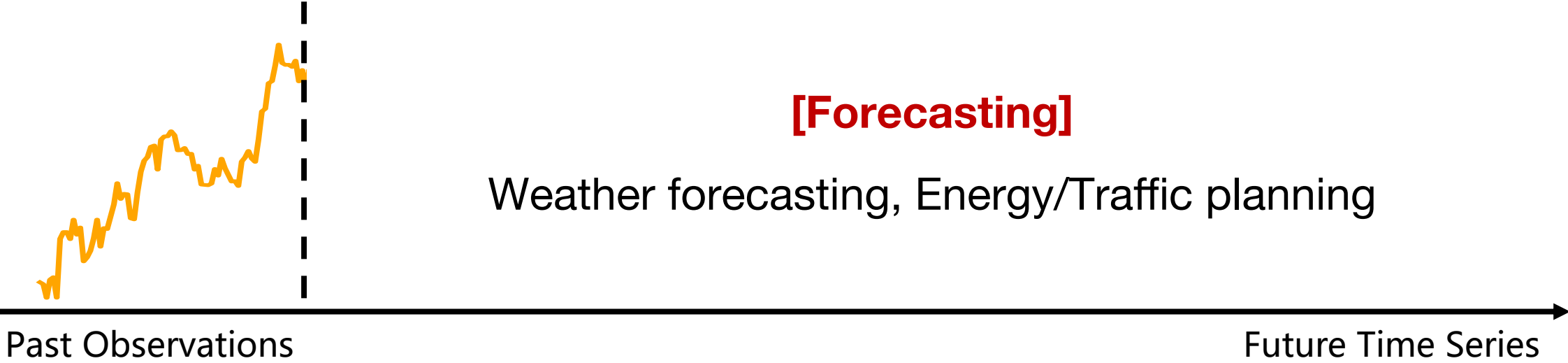


Time

Time Series Analysis



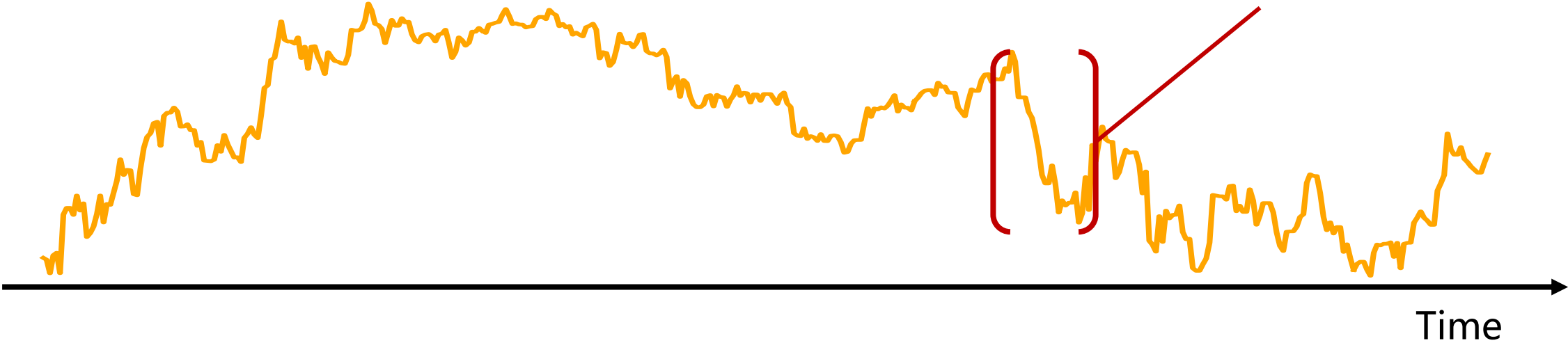
Time Series Analysis



Time Series Analysis

[Anomaly Detection]

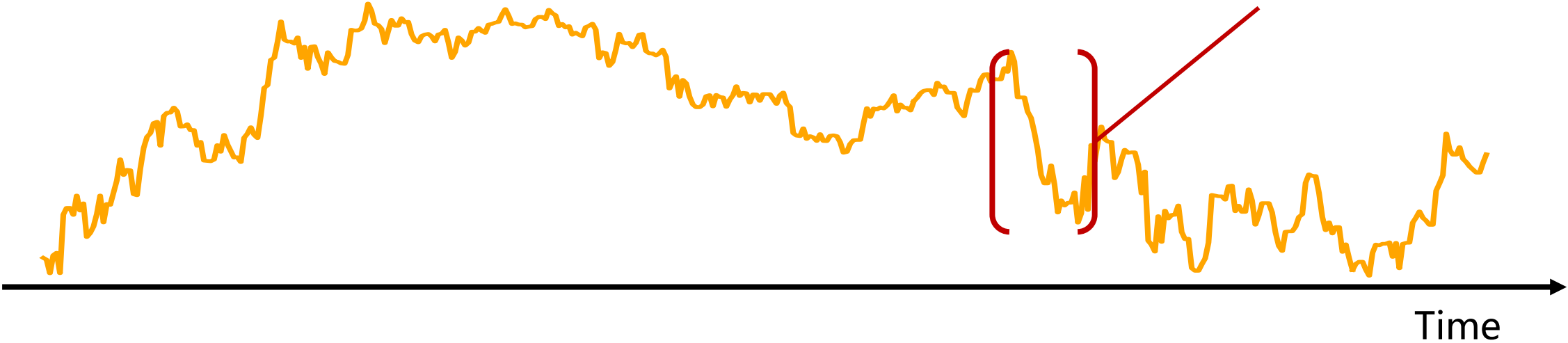
Industrial Maintenance



Time Series Analysis

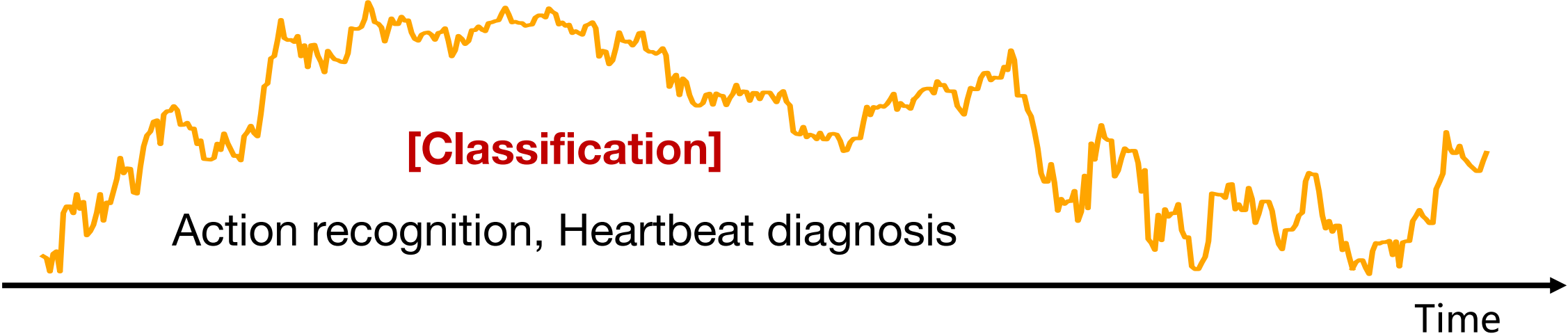
[Anomaly Detection]

Industrial Maintenance

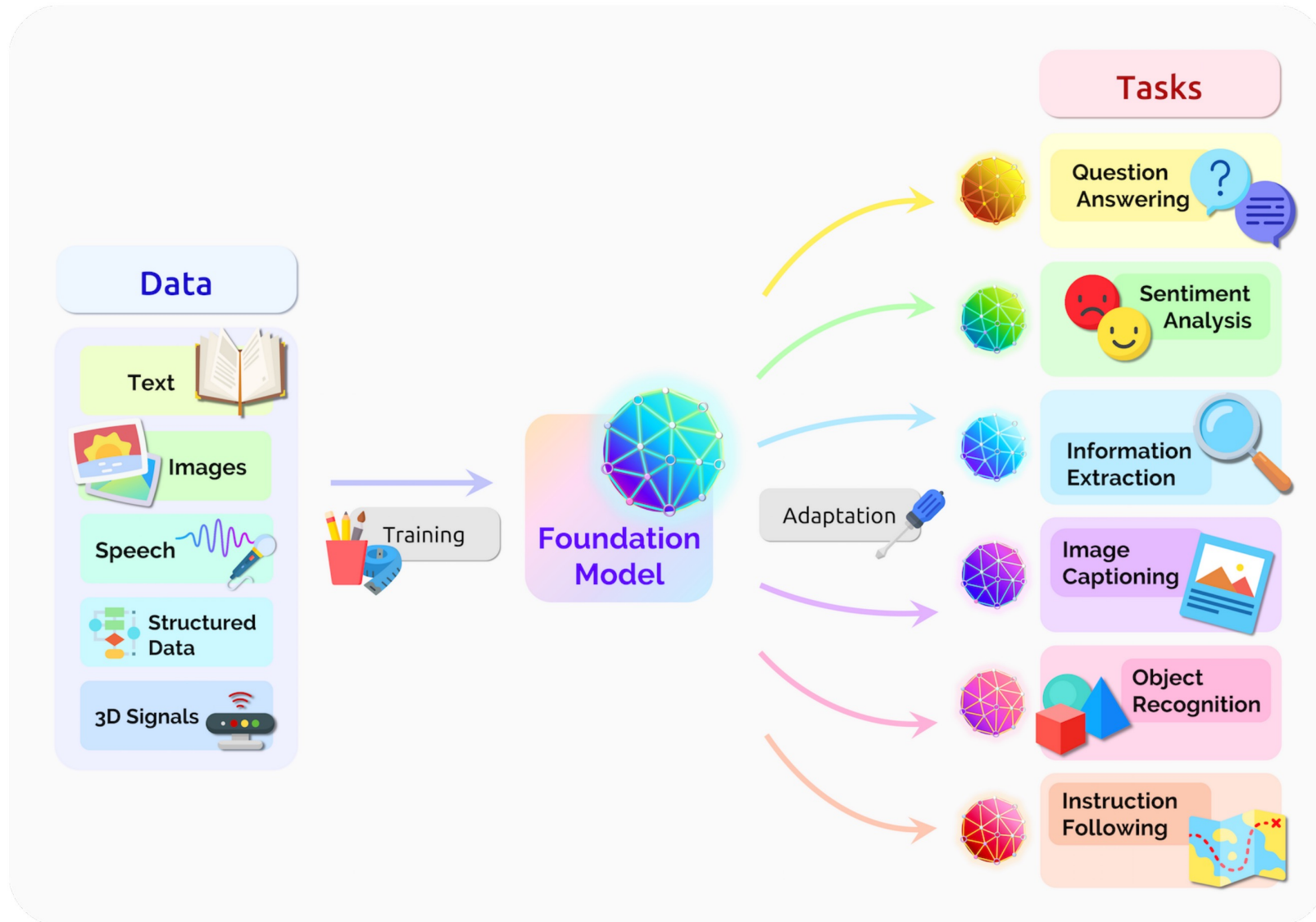


[Classification]

Action recognition, Heartbeat diagnosis



In Pursuing Foundation Models



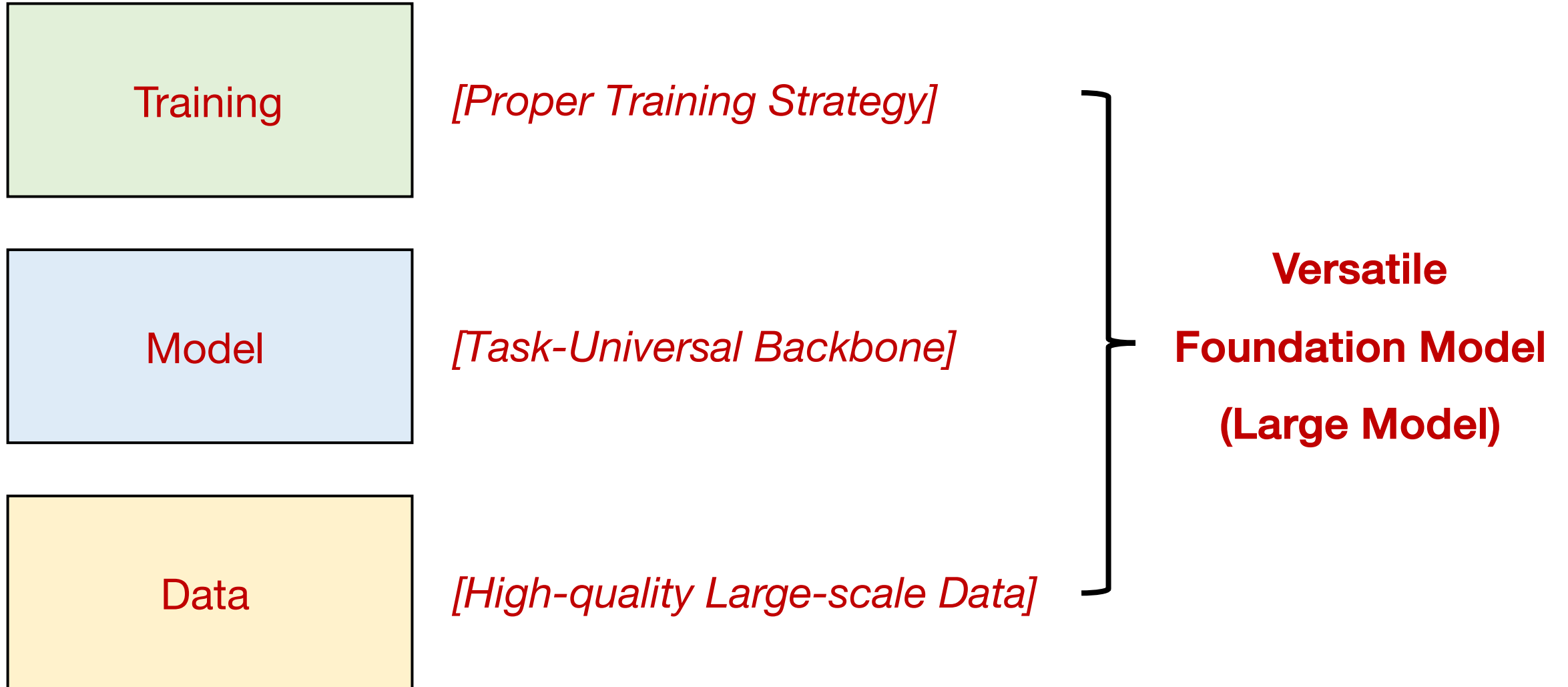
[Data Universal]

Learn from various modalities

[Task Universal]

Adapt to a wide range of
downstream tasks

In Pursuing Foundation Models



37th Conference on Neural Information Processing Systems (NeurIPS 2023).



SimMTM: A Simple Pre-Training Framework for Masked Time-Series Modeling

Jiaxiang Dong*, Haixu Wu*, Haoran Zhang, Li Zhang, Jianmin Wang, Mingsheng Long[✉]

School of Software, BNRist, Tsinghua University, China

{djx20, whx20, z-hr20}@mails.tsinghua.edu.cn, {lizhang, jimwang, mingsheng}@tsinghua.edu.cn



Jiaxiang Dong



Haixu Wu



Haoran Zhang



Li Zhang

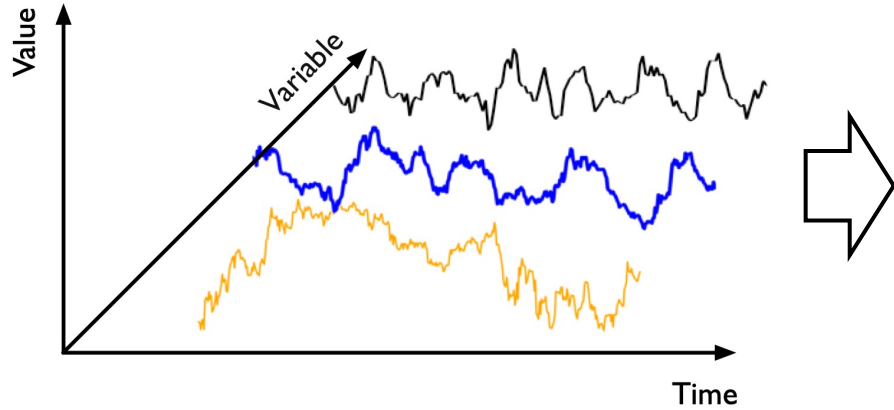


Jianmin Wang



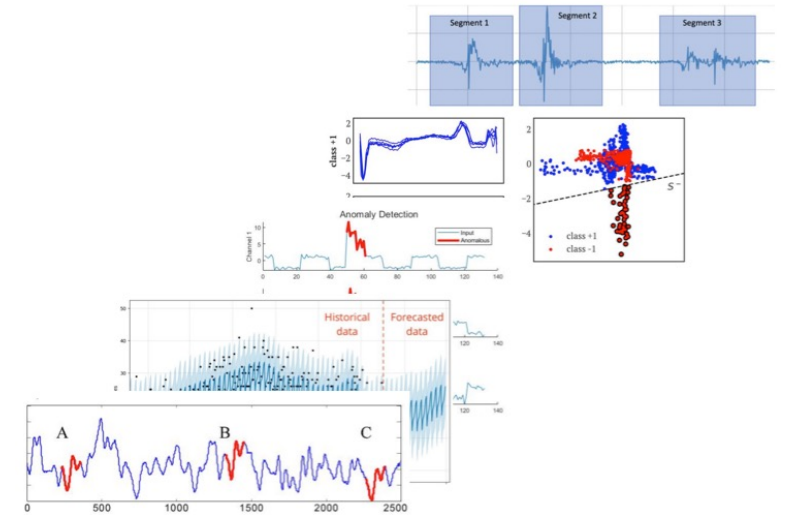
Mingsheng Long

Time Series Pre-training



Pre-training

Fine-tuning



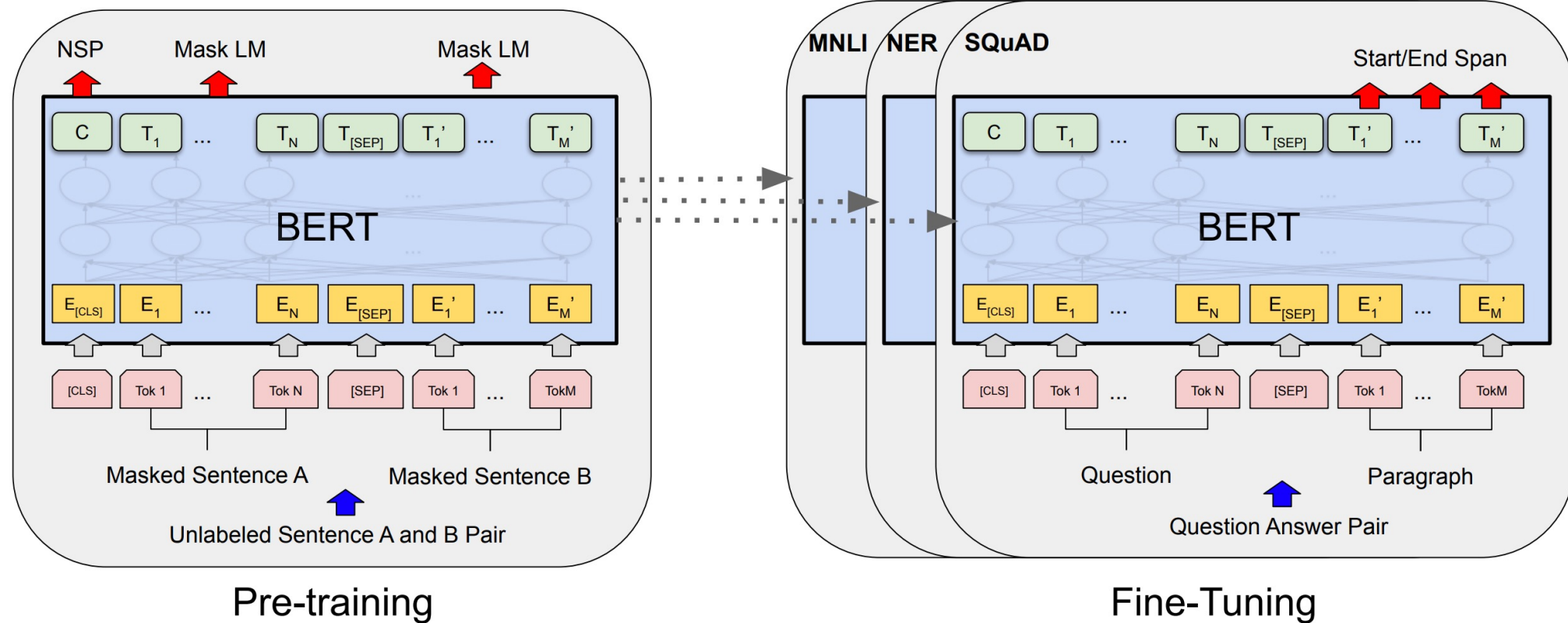
Large-scale time series data

Diversified time series analysis tasks

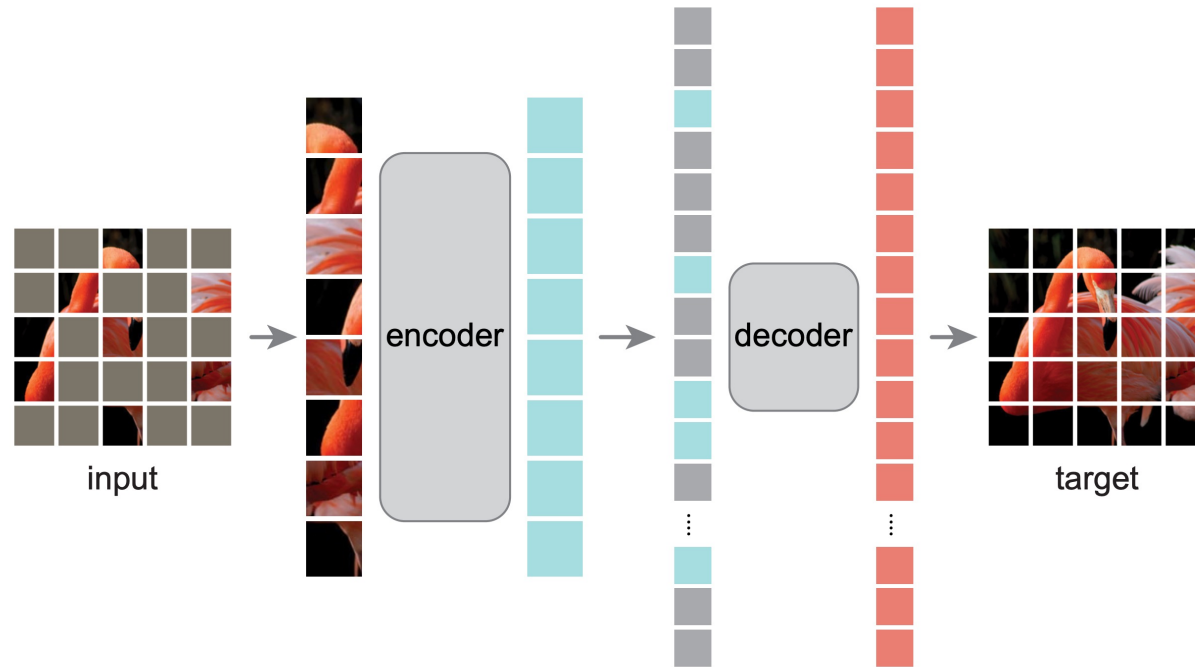
- ① Use the model as the carrier of knowledge.
- ② Learn transferable temporal representations.

Masked Modeling in NLP

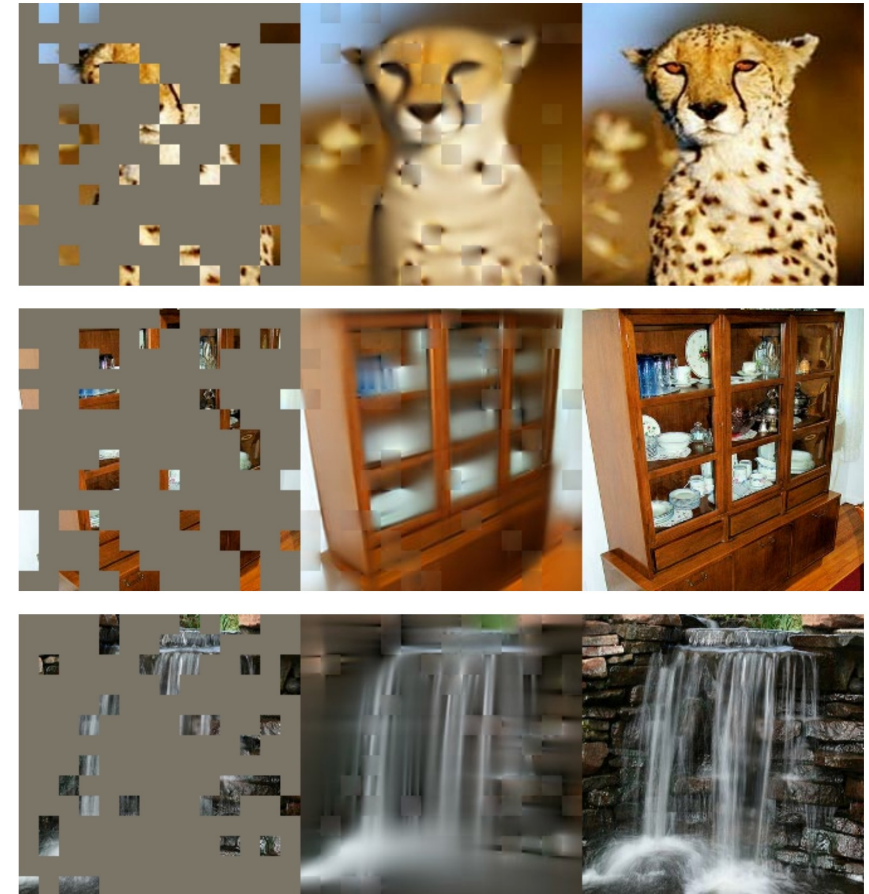
Random mask a portion of words.



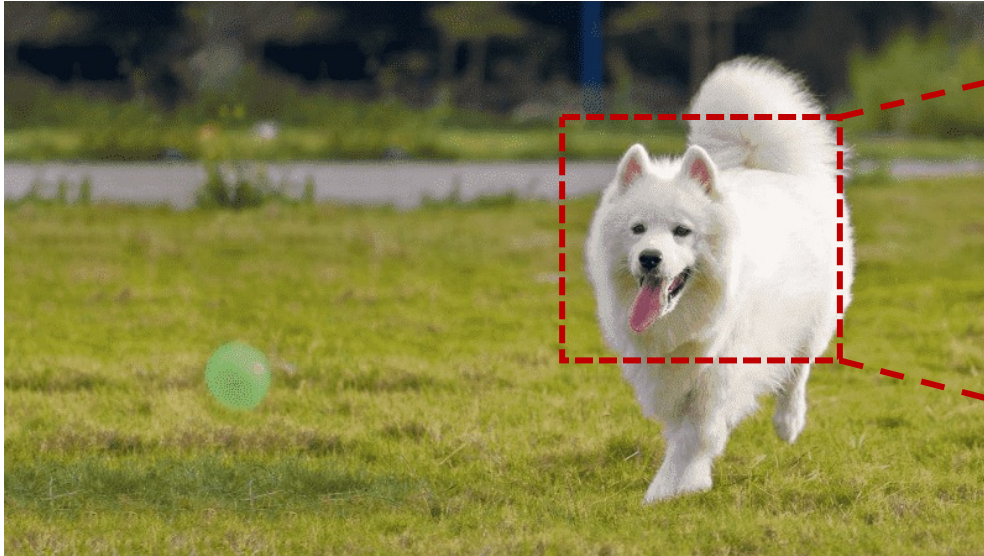
Masked Modeling in CV



Random mask a portion of patches.

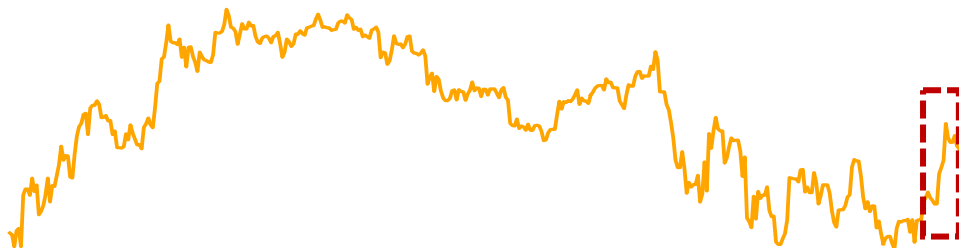


Differences among Image, Language, Time Series



TimesNet is for time series analysis.

Analysis is the process of breaking a complex topic into smaller parts for a better understanding.

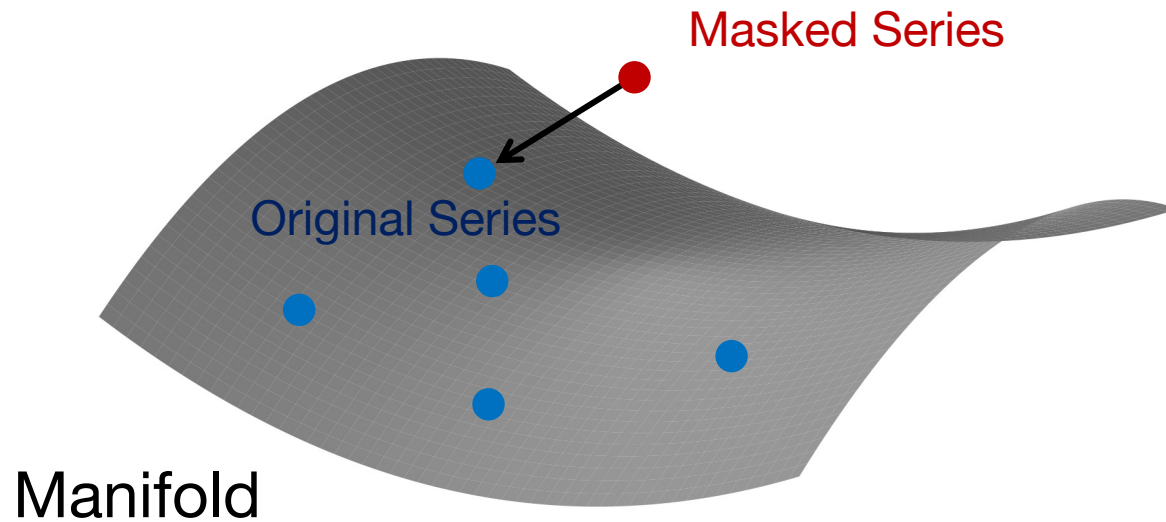


Each time point only saves some scalars.

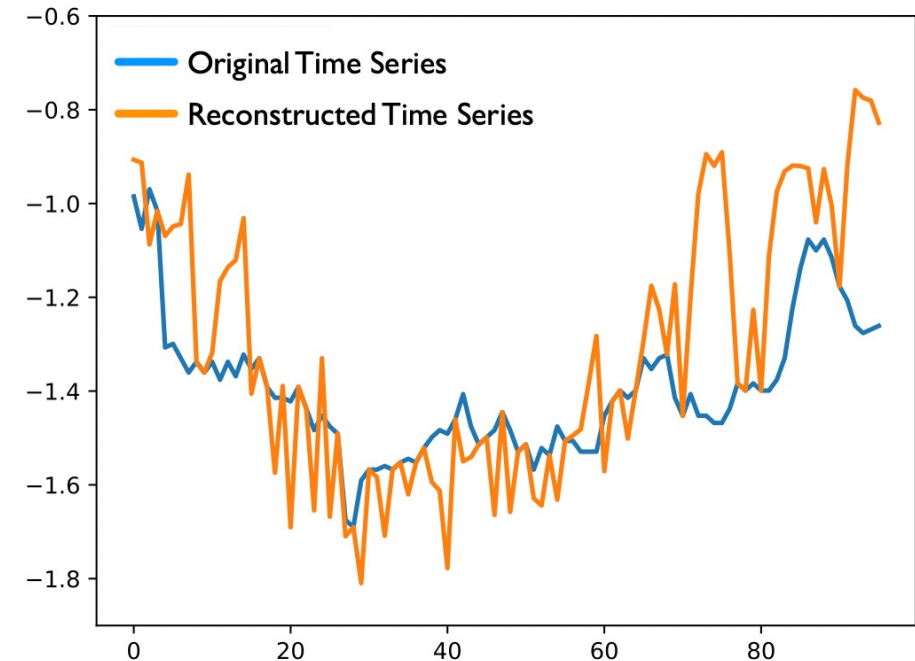


WIKIPEDIA
The Free Encyclopedia

Canonical Masked Modeling in Time Series



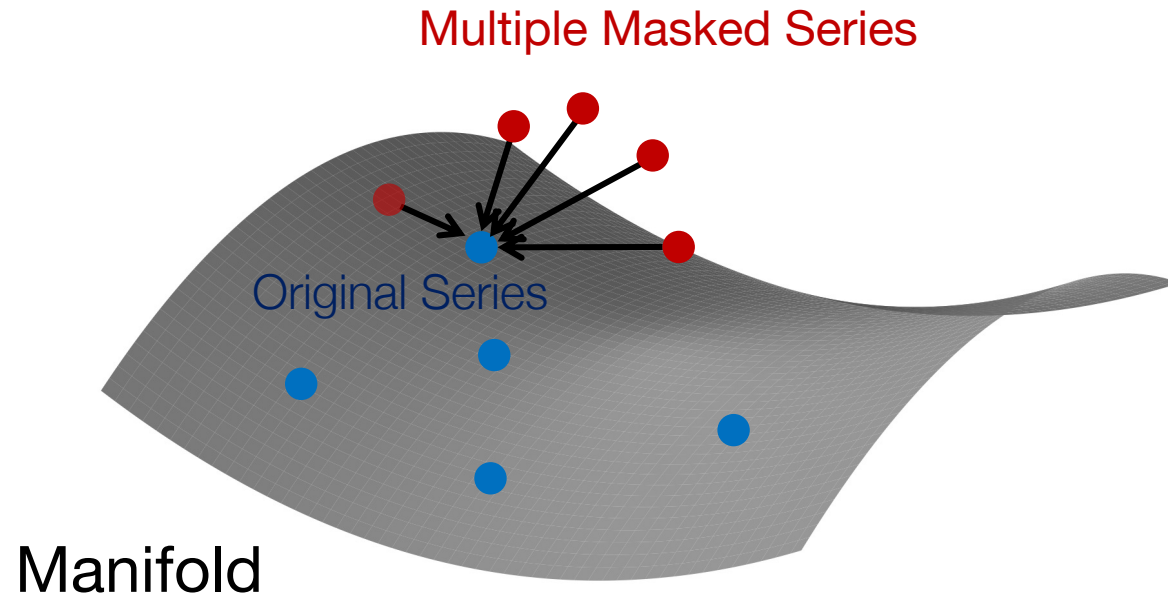
Hard to Reconstruct



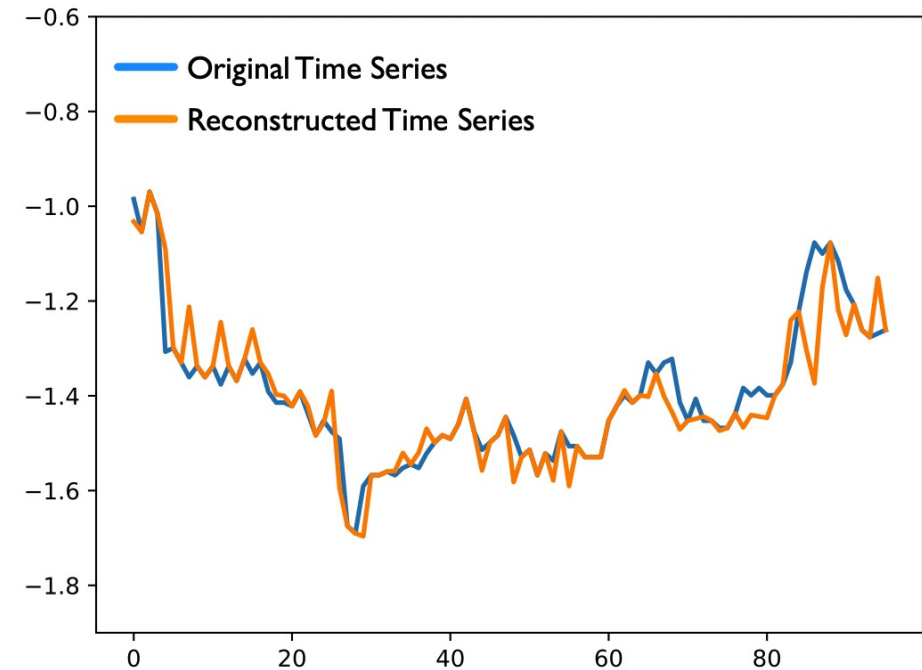
✓ Direct Reconstruction

Directly masking a portion of time points will seriously ruin the temporal variations of the original time series.

Multiple Masked Modeling



Benefit Masked Modeling

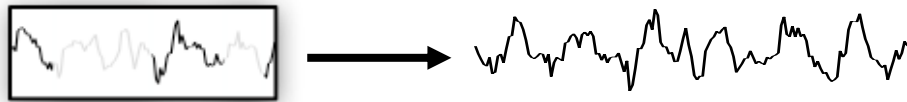


✓ Neighborhood Aggregation

Multiple randomly masked series will complement each other.

Neighborhood Aggregation Masked Modeling

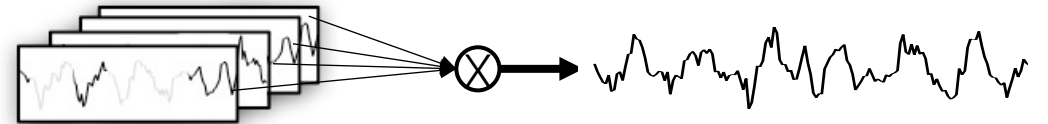
Canonical



- ✗ Critical information destruction
- ✗ Mask ratio sensitive
- ✗ Reconstruction difficulty

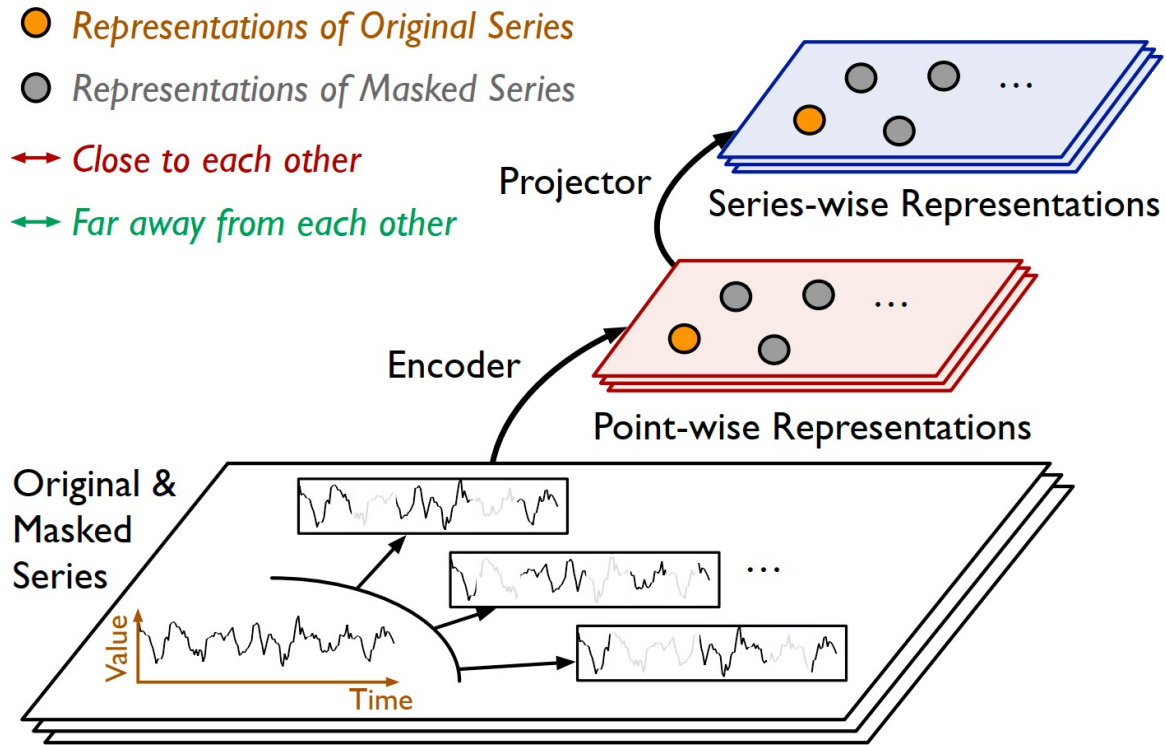
VS

Neighborhood Aggregation



- ✓ Multi-information perspective
- ✓ Information complementation
- ✓ Learnable aggregate weight

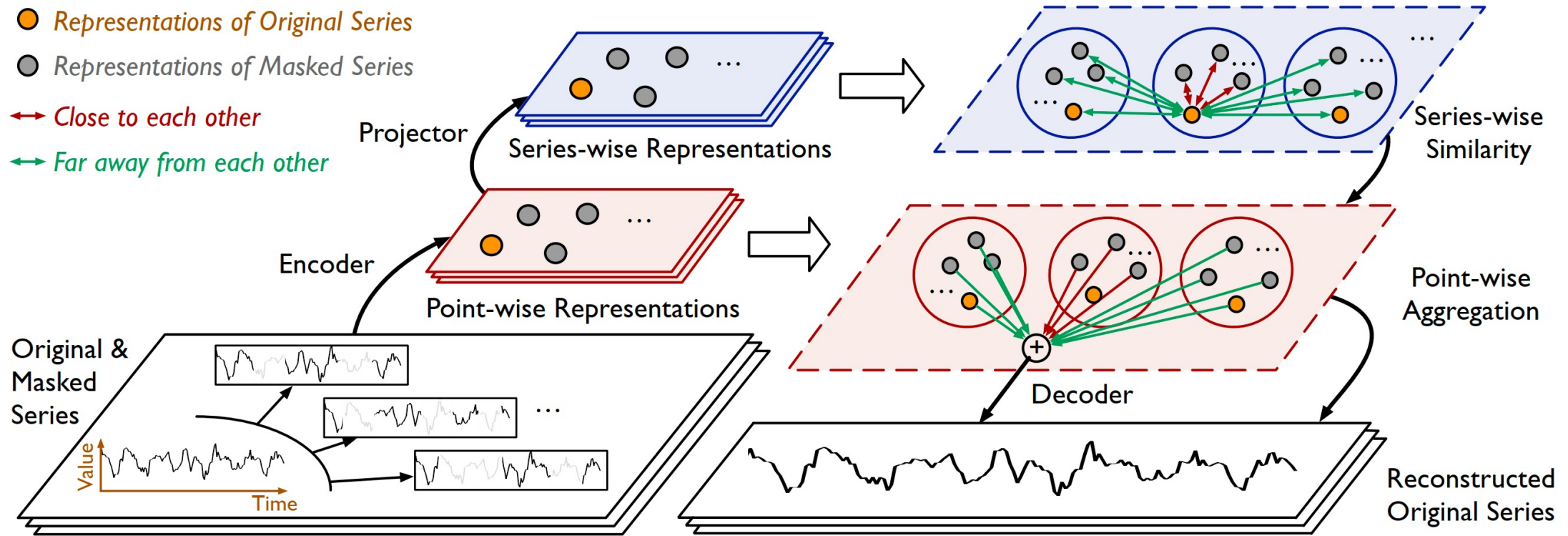
Overall design of SimMTM



Generate original & masked series representations.

- ① Point-wise Representations
- ② Series-wise Representations

Overall design of SimMTM



① Series-wise Similarity ② Point-wise Aggregation

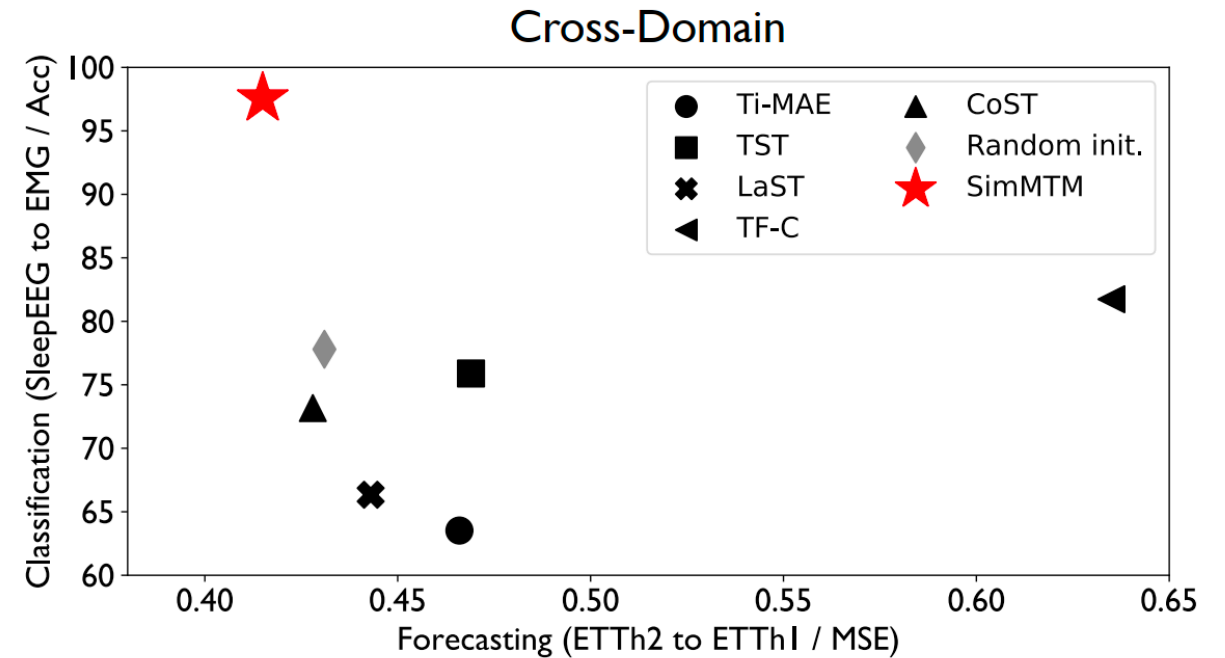
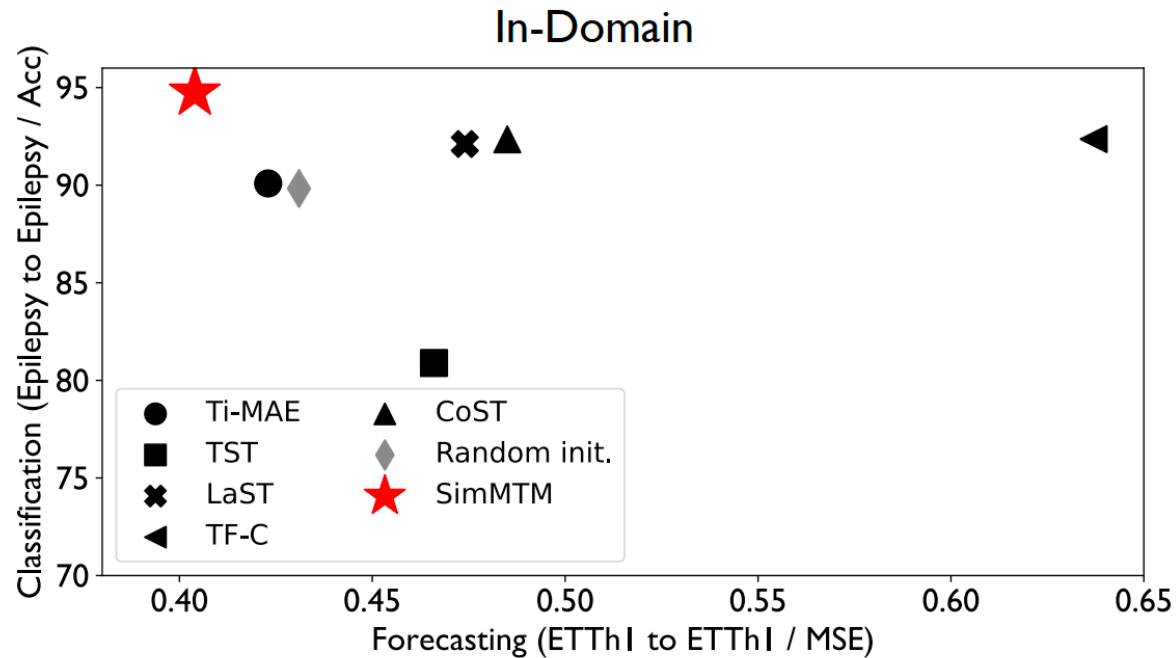
Multiple masked series complete each other and adaptive aggregate weight.

Experiment: Overall

Tasks	Datasets	Semantic
Forecasting	ETTh1,ETTh2	Electricity
	ETTm1,ETTm2	Electricity
	Weather	Weather
	Electricity	Electricity
	Traffic	Transportation
Classification	SleepEEG	EEG
	Epilepsy	EEG
	FD-B	Faulty Detection
	Gesture	Hand Movement
	EMG	Muscle Responses

- ✓ Two typical time series analysis tasks: **Forecasting and Classification.**
- ✓ Under multiple experiment settings: **In- and Cross domain**
- ✓ Compared to **6 advanced baselines in 12 databases.**

Experiment: Overall



SimMTM pretraining can benefit

both forecasting and classification tasks.

Model Generality on diverse base models

Dataset	ETTh1		ETTh2		ETTm1		ETTm2	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Transformer [39]	1.088	0.836	4.103	1.612	0.901	0.704	1.624	0.901
+ SimMTM	0.927	0.761	3.498	1.487	0.809	0.663	1.322	0.808
Autoformer [47]	0.573	0.573	0.550	0.559	0.615	0.528	0.324	0.368
+ SimMTM	0.561	0.568	0.543	0.555	0.553	0.505	0.315	0.360
NS Transformer [24]	0.570	0.537	0.526	0.516	0.481	0.456	0.306	0.347
+ SimMTM	0.543	0.527	0.493	0.514	0.431	0.455	0.301	0.345
PatchTST [26]	0.417	0.431	0.331	0.379	0.352	0.382	0.258	0.317
+ Sub-series Masking	0.430↓	0.445↓	0.355↓	0.394↓	0.341	0.379	0.258	0.318↓
+ SimMTM	0.409	0.428	0.329	0.379	0.348	0.378	0.254	0.313



SimMTM can consistently improve
the forecasting performance of **diverse base models.**

Open Source

thuml / SimMTM

Code Issues 1 Pull requests Actions Projects Wiki Security Insights Settings

SimMTM Public

main 1 Branch 0 Tags

Go to file Add file Code

dongjiaxiang adjust readme b4b676e · last month 2 Commits

- SimMTM_Classification init commit last month
- SimMTM_Forecasting init commit last month
- figs init commit last month
- README.md adjust readme last month

SimMTM (NeurIPS 2023)

This is the codebase for the paper: [SimMTM: A Simple Pre-Training Framework for Masked Time-Series Modeling](#)

Architecture

Figure 1. Overview of SimMTM.

The reconstruction process of SimMTM involves the following four modules: masking, representation learning, series-wise similarity learning and point-wise reconstruction.

About

About Code release for "SimMTM: A Simple Pre-Training Framework for Masked Time-Series Modeling" (NeurIPS 2023 Spotlight), <https://arxiv.org/abs/2302.00861>

Releases

Package

Languages

- Python 96.9%
- Shell 3.1%

Suggested workflows

Python package Configure

Code is available at <https://github.com/thuml/SimMTM>



Published as a conference paper at ICLR 2023

TIMESNET: TEMPORAL 2D-VARIATION MODELING FOR GENERAL TIME SERIES ANALYSIS

Haixu Wu*, **Tengge Hu***, **Yong Liu***, **Hang Zhou**, **Jianmin Wang**, **Mingsheng Long**✉

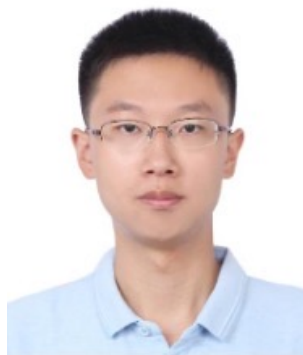
School of Software, BNRist, Tsinghua University, Beijing 100084, China

{whx20, liuyong21, htg21, h-zhou18}@mails.tsinghua.edu.cn

{jimwang, mingsheng}@tsinghua.edu.cn



Haixu Wu



Tengge Hu



Yong Liu



Hang Zhou



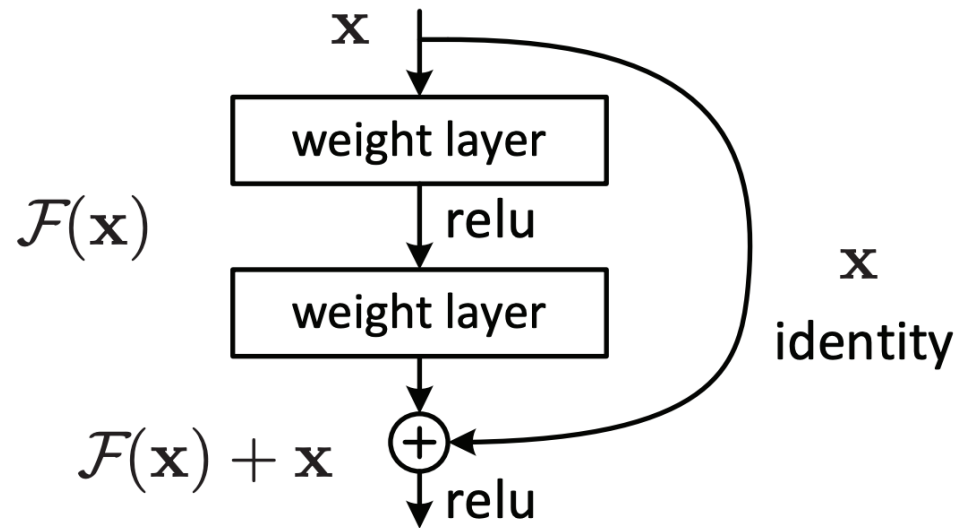
Jianmin Wang



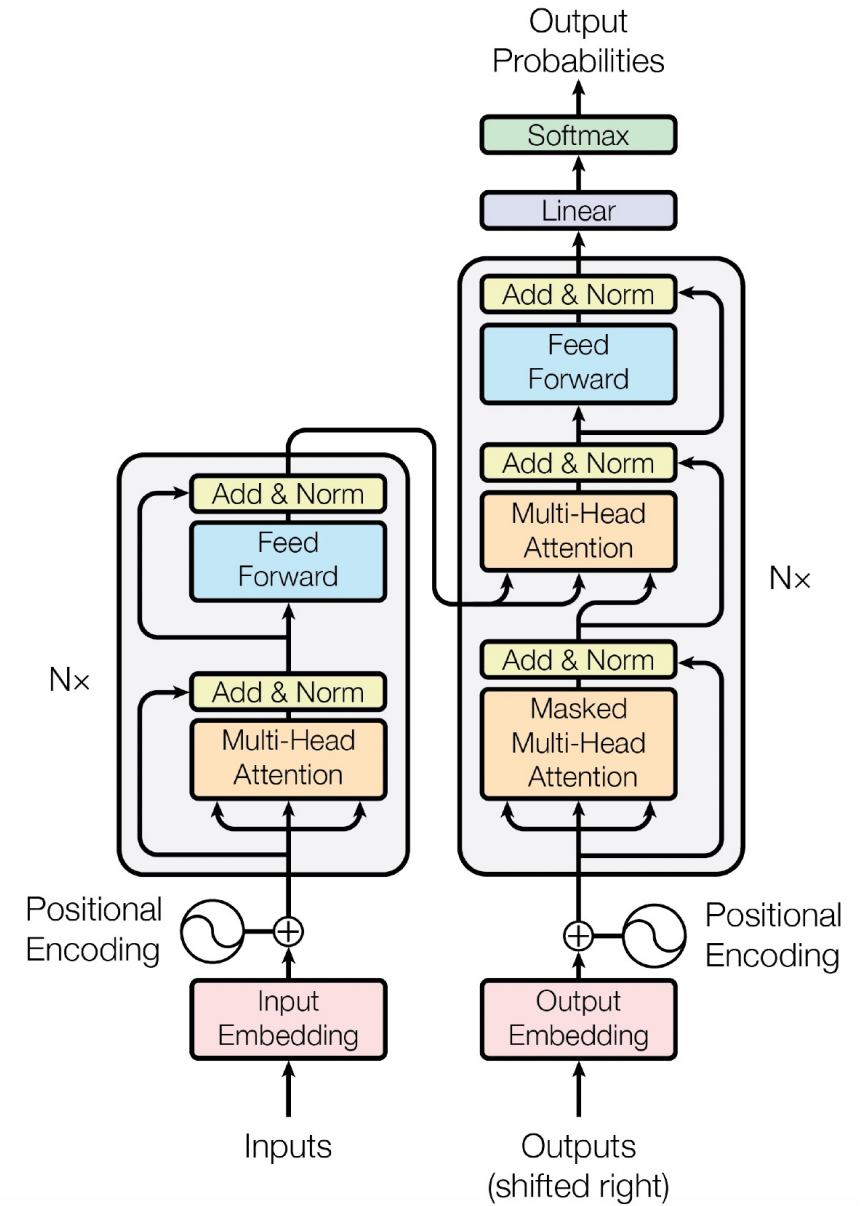
Mingsheng Long

Foundation Models in CV and NLP

Universal backbone with **task-specific heads** for different tasks.



Classification, Object detection, Segmentation

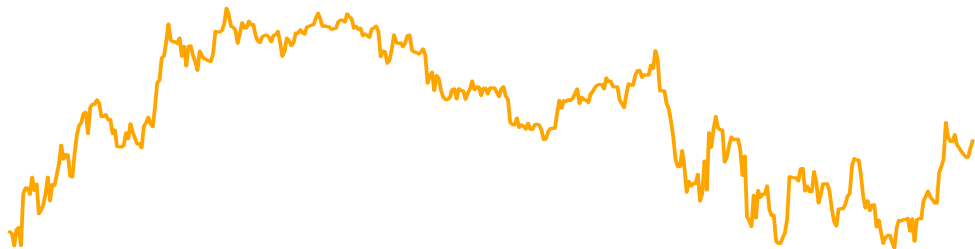


Classification, Generation

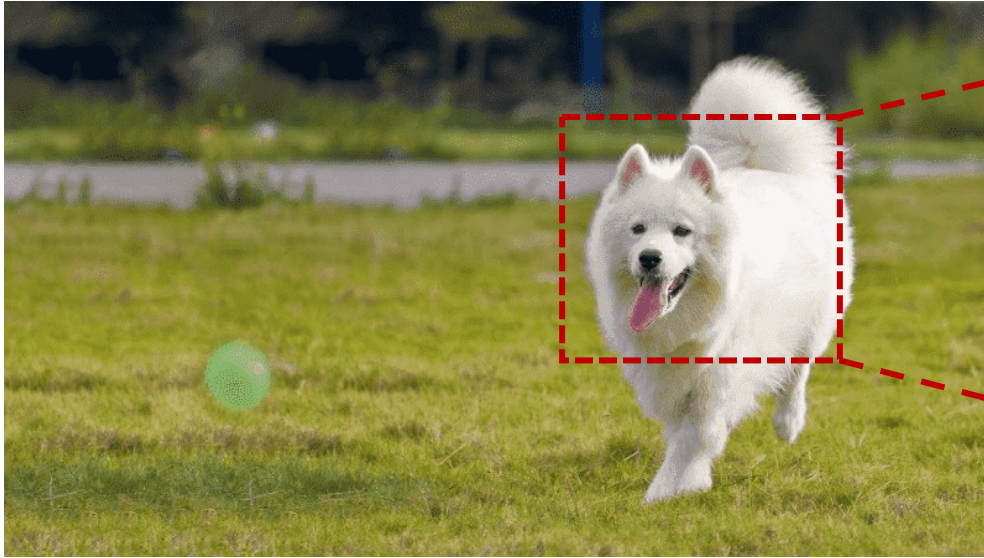
Differences among Image, Language, Time Series



TimesNet is for time series analysis.

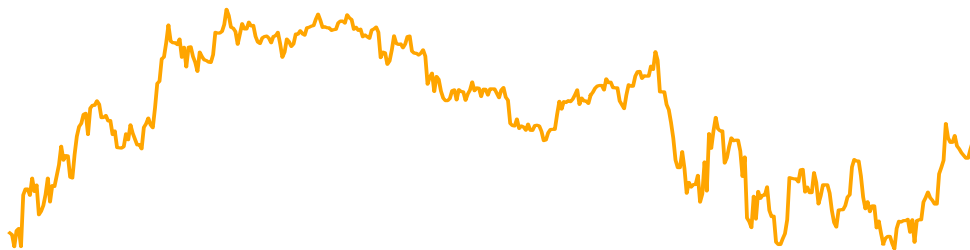


Differences among Image, Language, Time Series



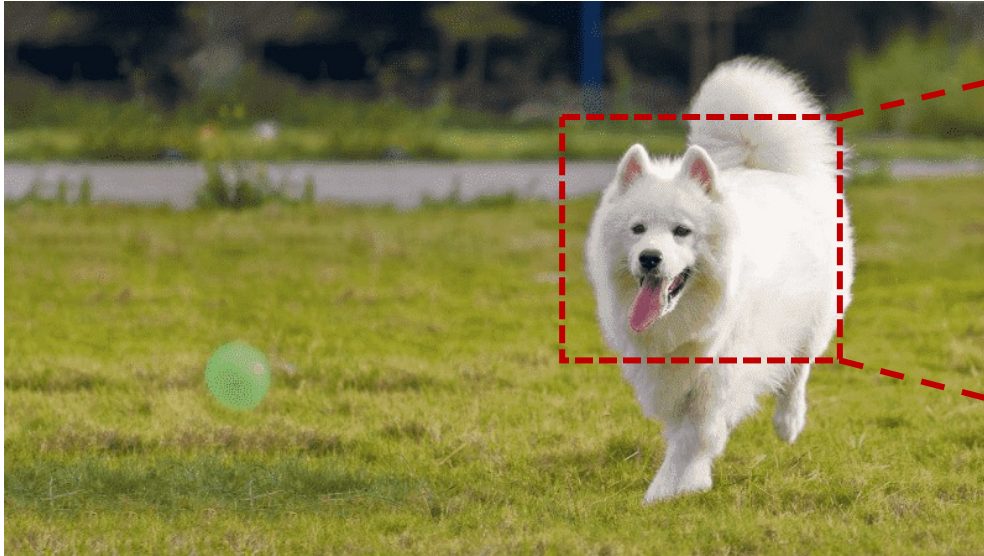
TimesNet is for time series analysis.

Analysis is the process of breaking a complex topic into smaller parts for a better understanding.



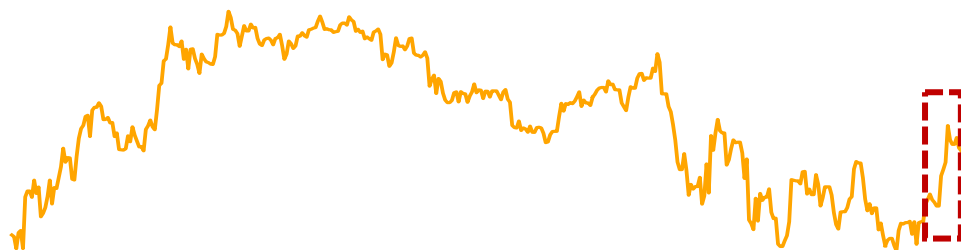
WIKIPEDIA
The Free Encyclopedia

Differences among Image, Language, Time Series



TimesNet is for time series analysis.

Analysis is the process of breaking a complex topic into smaller parts for a better understanding.



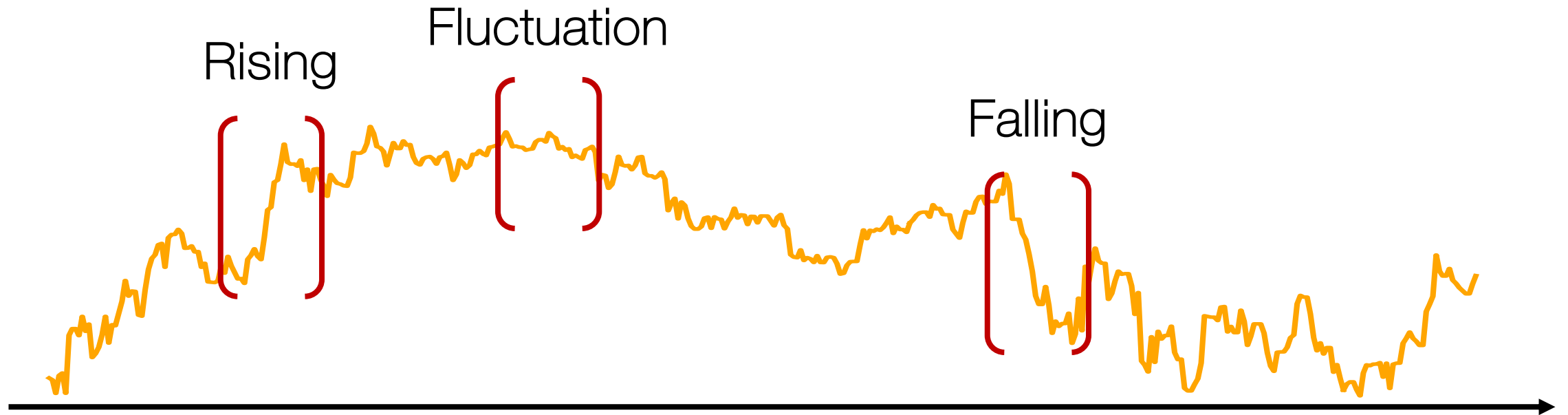
Each time point only saves some scalars.



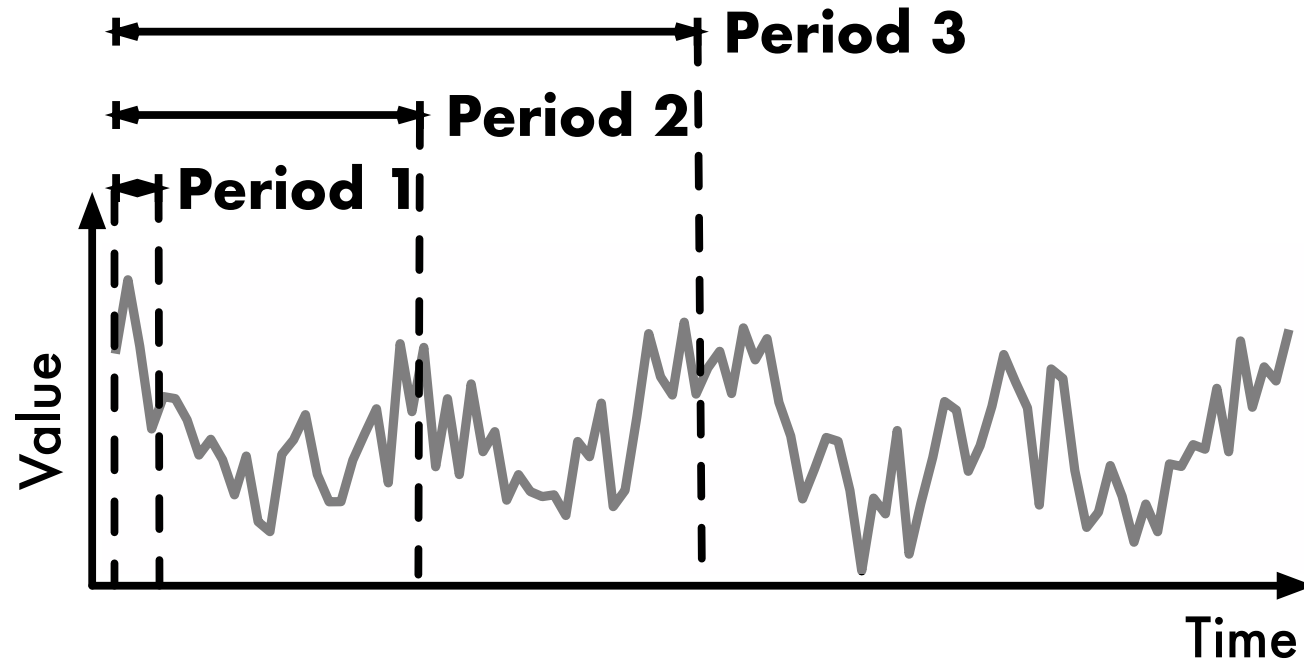
WIKIPEDIA
The Free Encyclopedia

Temporal Variations of Time Series

More information of time series is in **temporal variations**, such as continuity, periodicity, trend and etc.



Multi-periodicity View of Time Series

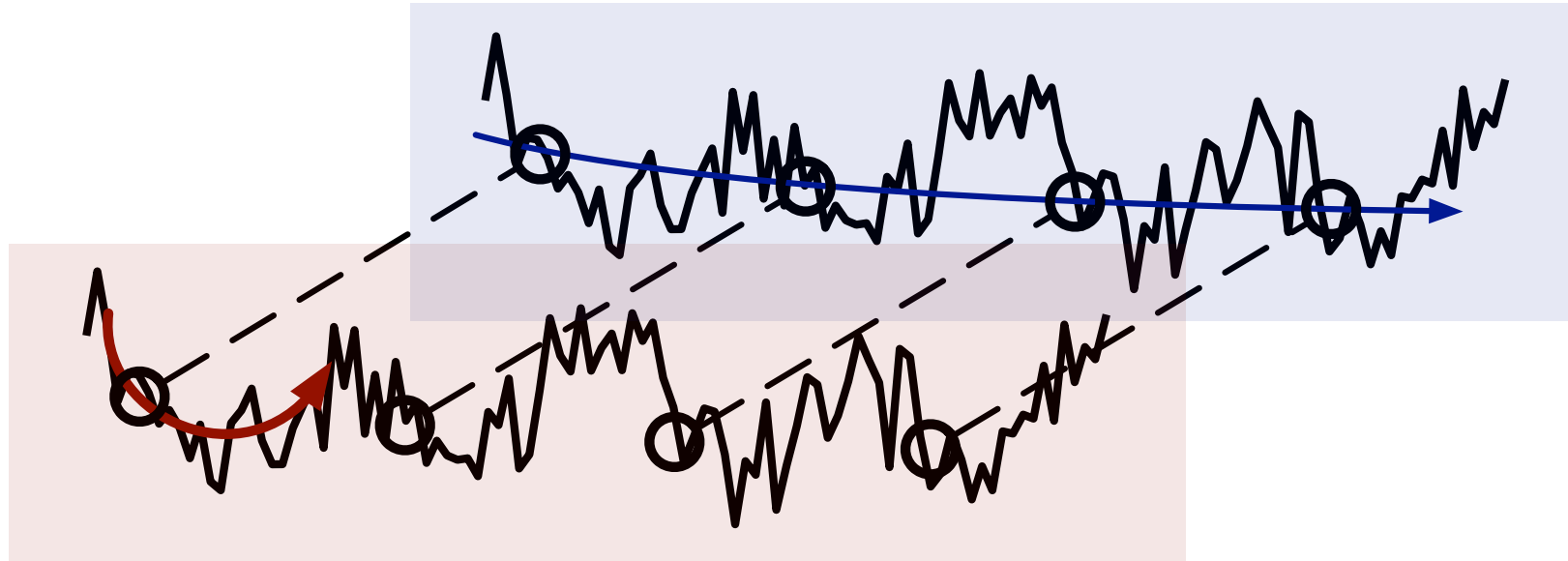


- ✓ Traffic: daily and weekly
- ✓ Weather: daily and yearly

Real-world time series usually present multi-periodicity.

Multiple periods overlap and interact with each other.

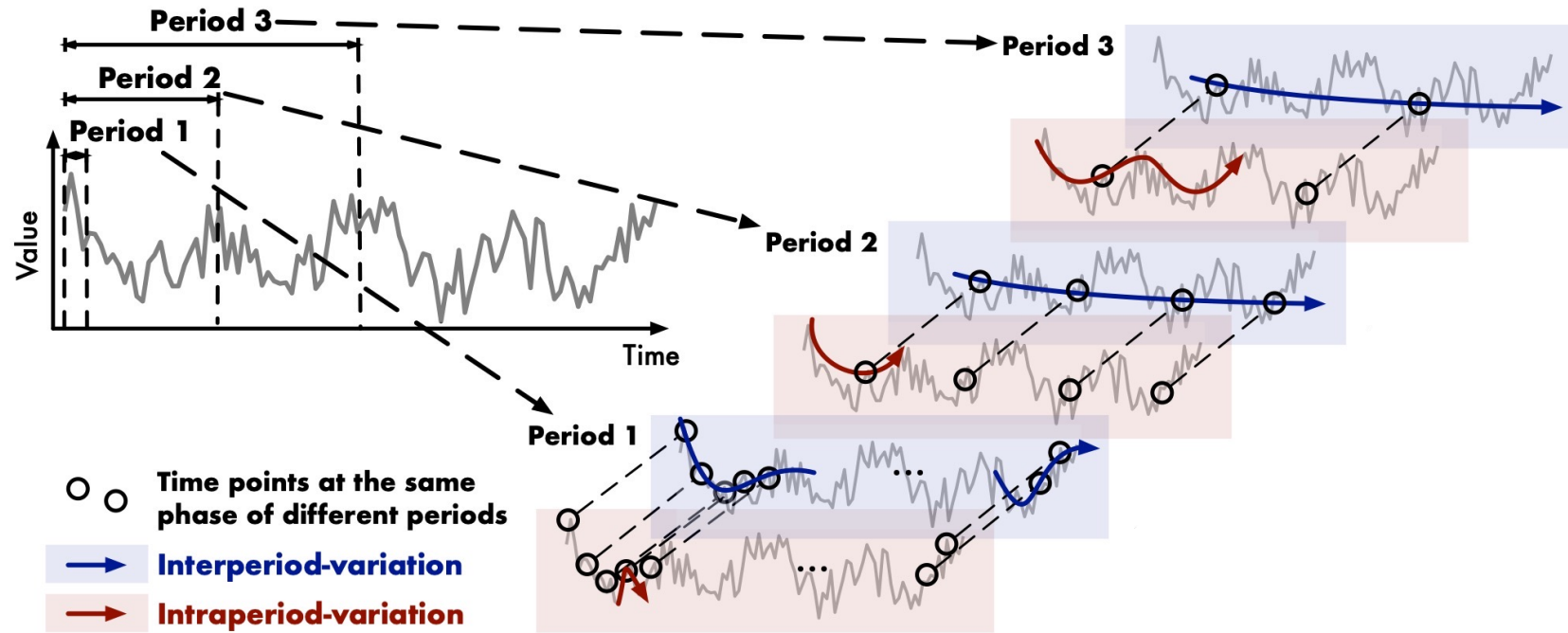
Intraperiod- and Interperiod-variations



- ✓ **Intraperiod**: adjacent area, **short-term variations**
- ✓ **Interperiod**: same phase in adjacent periods, **long-term variations**

Non-periodic cases, the variations will be dominated by intraperiod-variations.

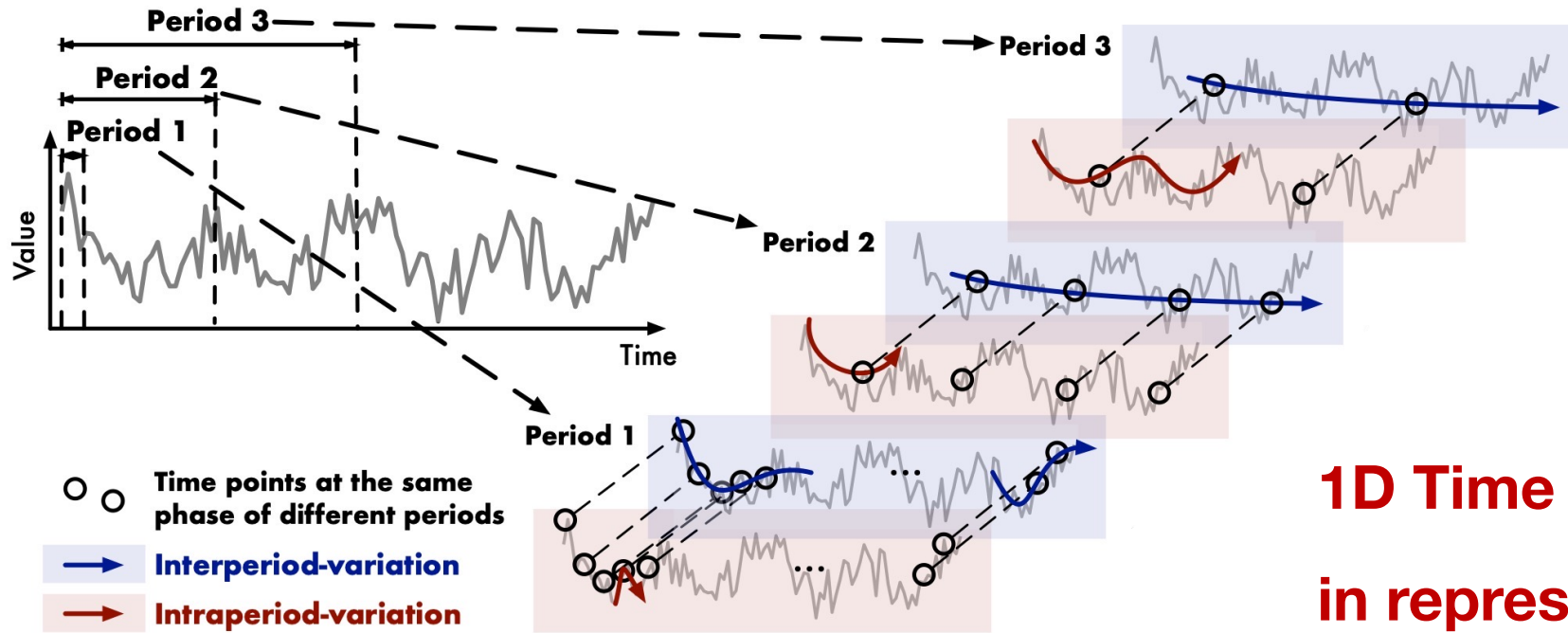
Overall design of TimesNet



① Multi-periodicity

A modular architecture to disentangle intricate temporal patterns

Overall design of TimesNet



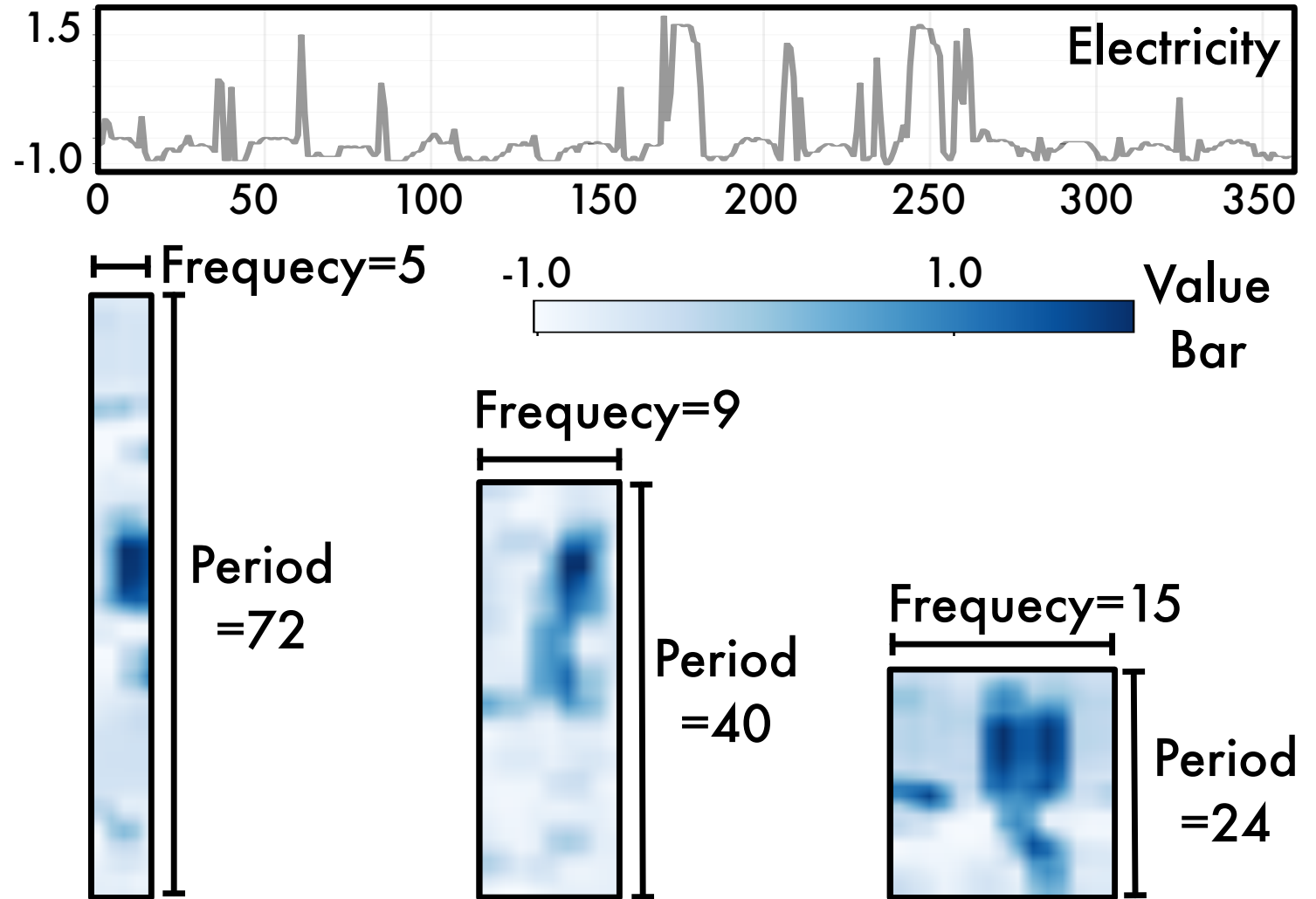
1D Time Series has limitations in representation capability.

① Multi-periodicity

A modular architecture to disentangle intricate temporal patterns

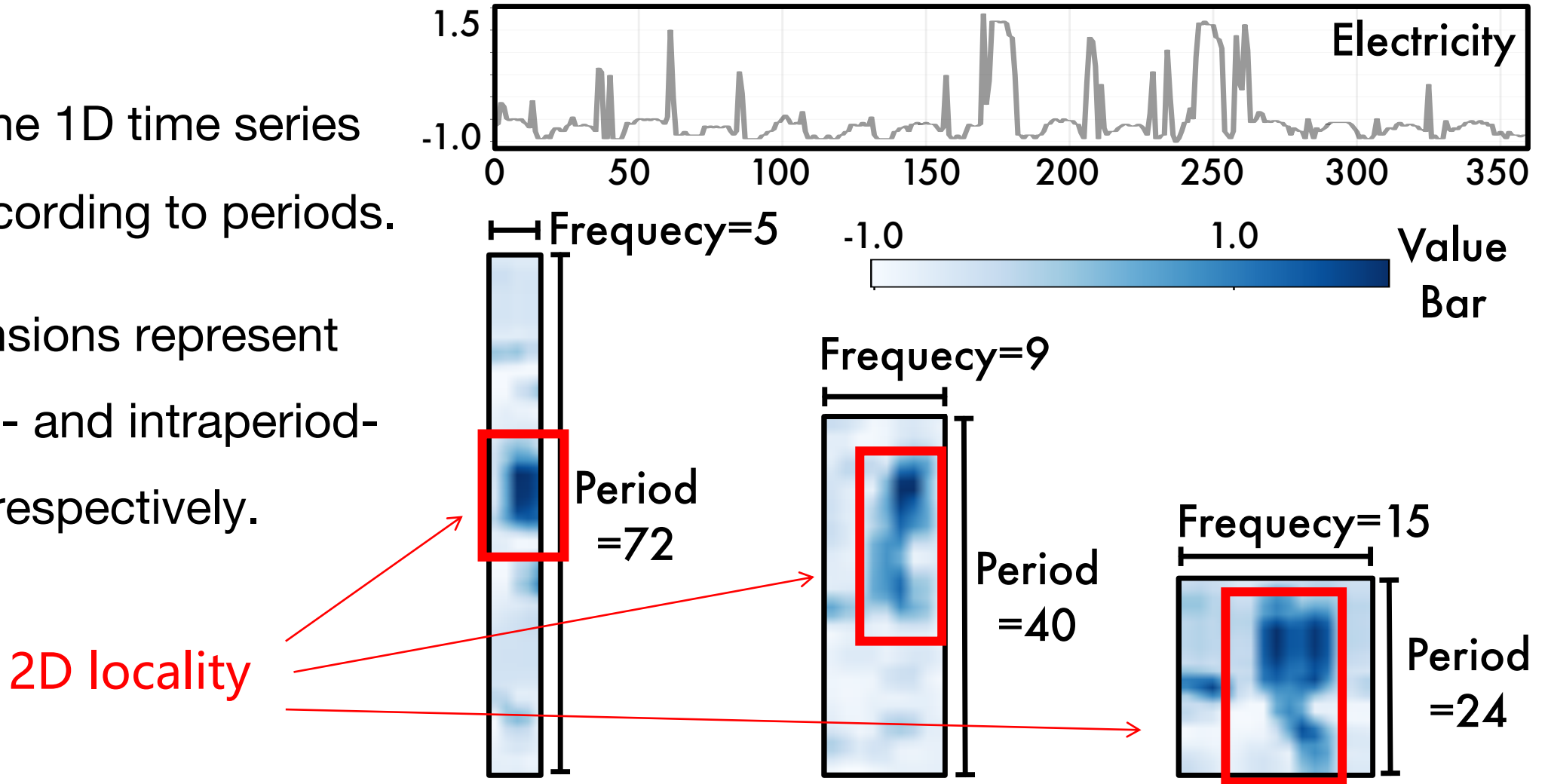
Temporal 2D-variation: A Case Study

- ✓ Reshape the 1D time series into 2D according to periods.
- ✓ Two dimensions represent interperiod- and intraperiod-variations respectively.



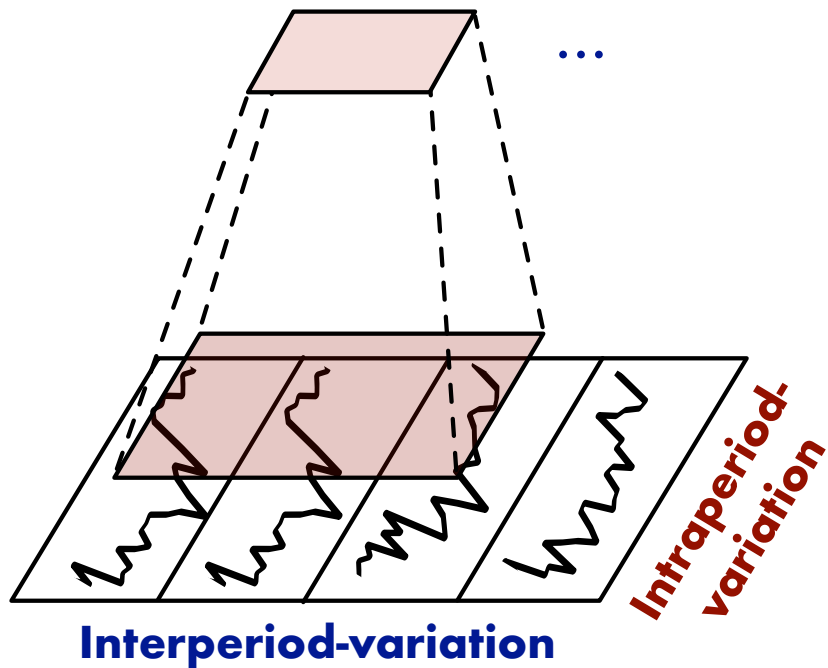
Temporal 2D-variation: A Case Study

- ✓ Reshape the 1D time series into 2D according to periods.
- ✓ Two dimensions represent interperiod- and intraperiod-variations respectively.



Temporal 2D-variation: A Case Study

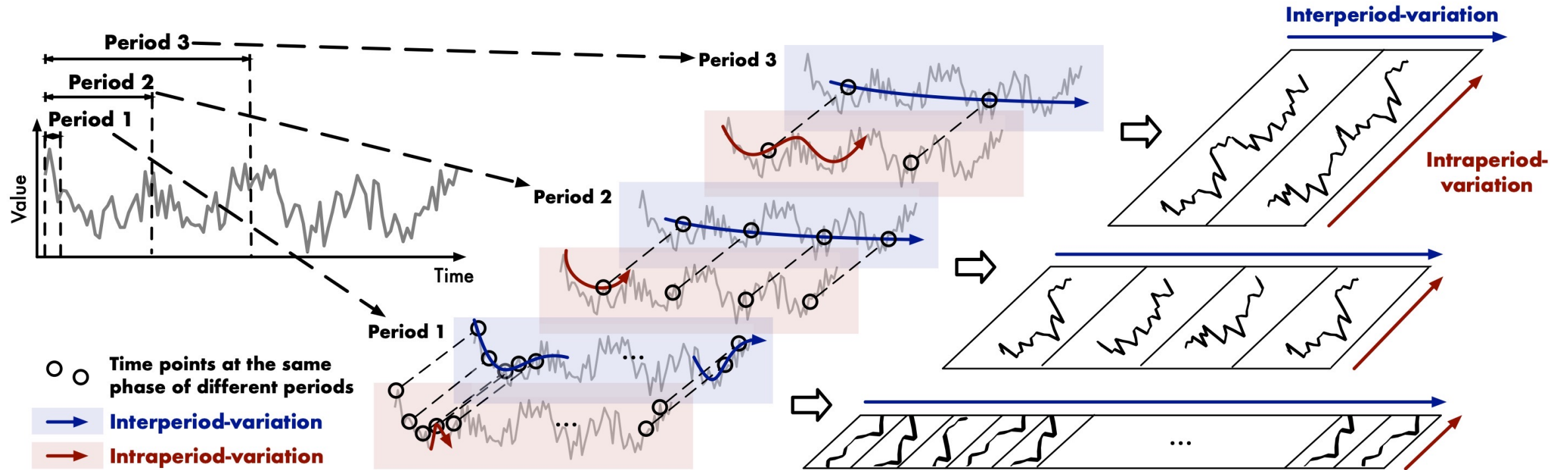
Capture Temporal 2D-variations
by 2D Kernels



With temporal 2D-variations, we can

- ✓ Unify intraperiod- interperiod-variations
- ✓ Learn representations by 2D kernels

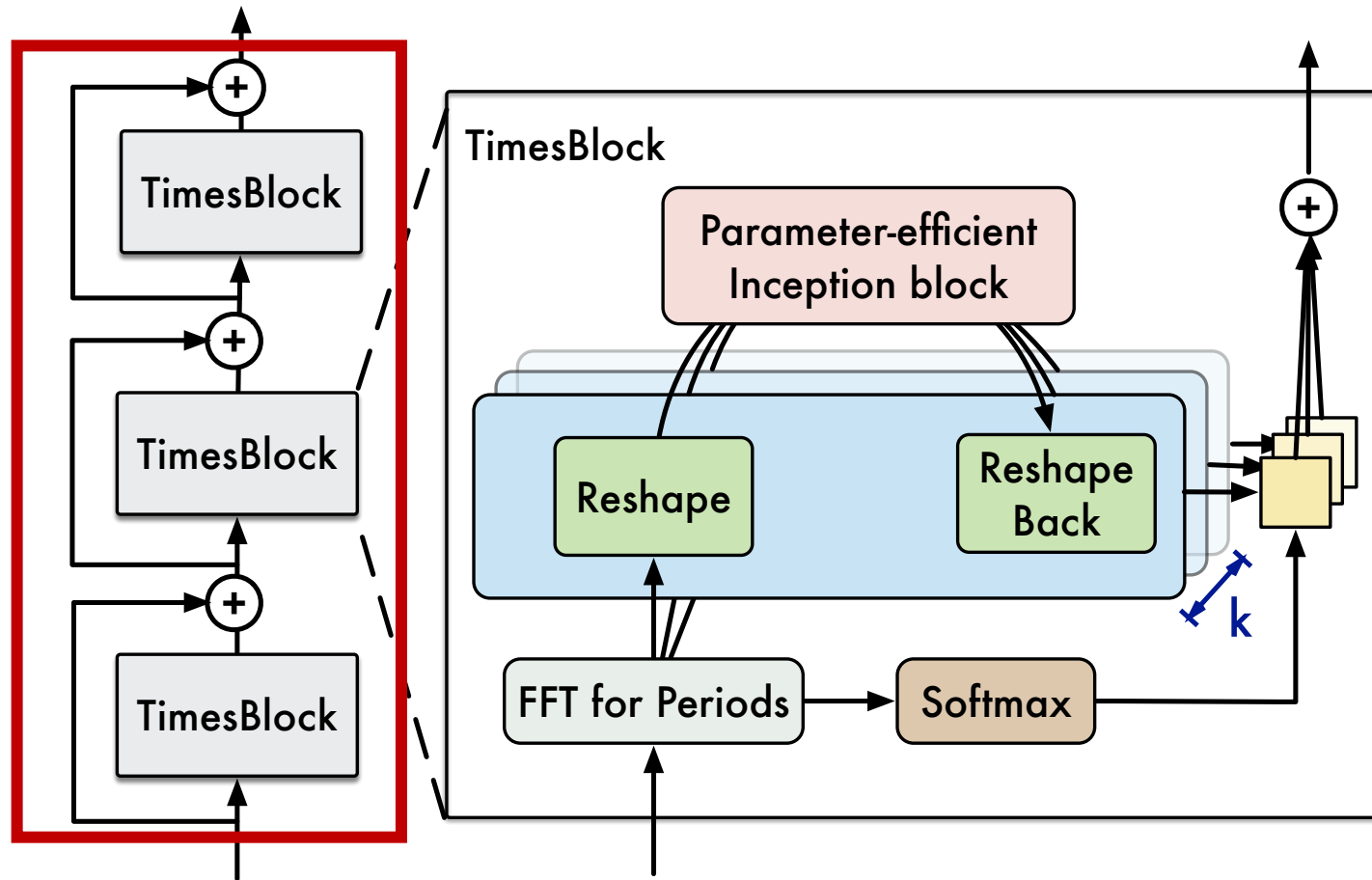
Overall design of TimesNet



① Multi-periodicity ② Temporal 2D-variation

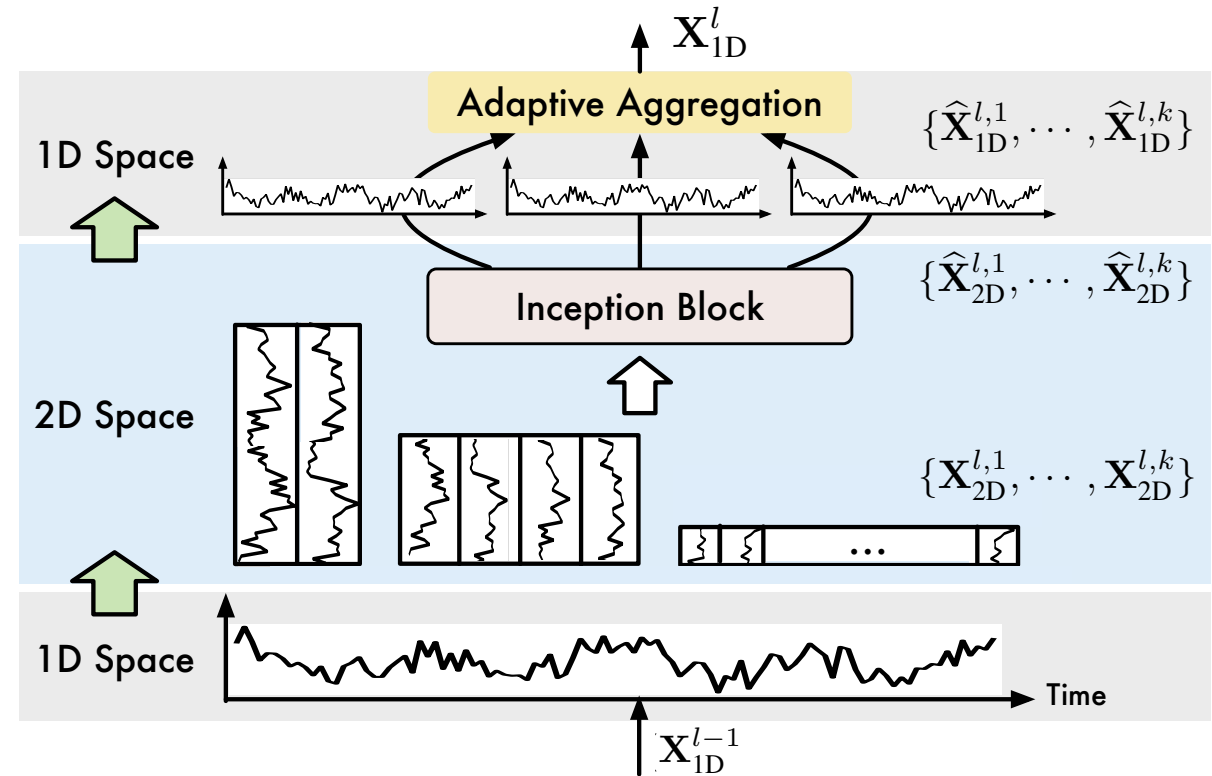
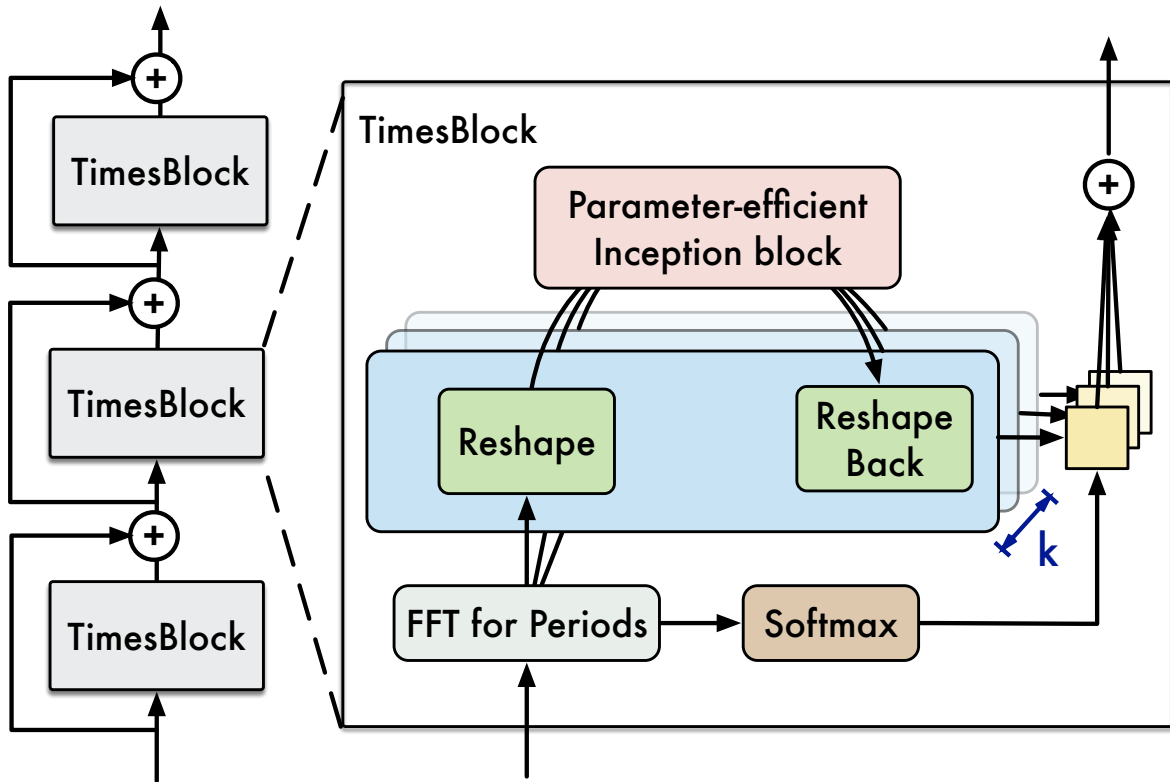
Unify intraperiod- and interperiod-variations in 2D

TimesNet



TimesNet consists of residual-connected **TimesBlocks**.

TimesBlock



TimesBlock learns representations in 2D space.

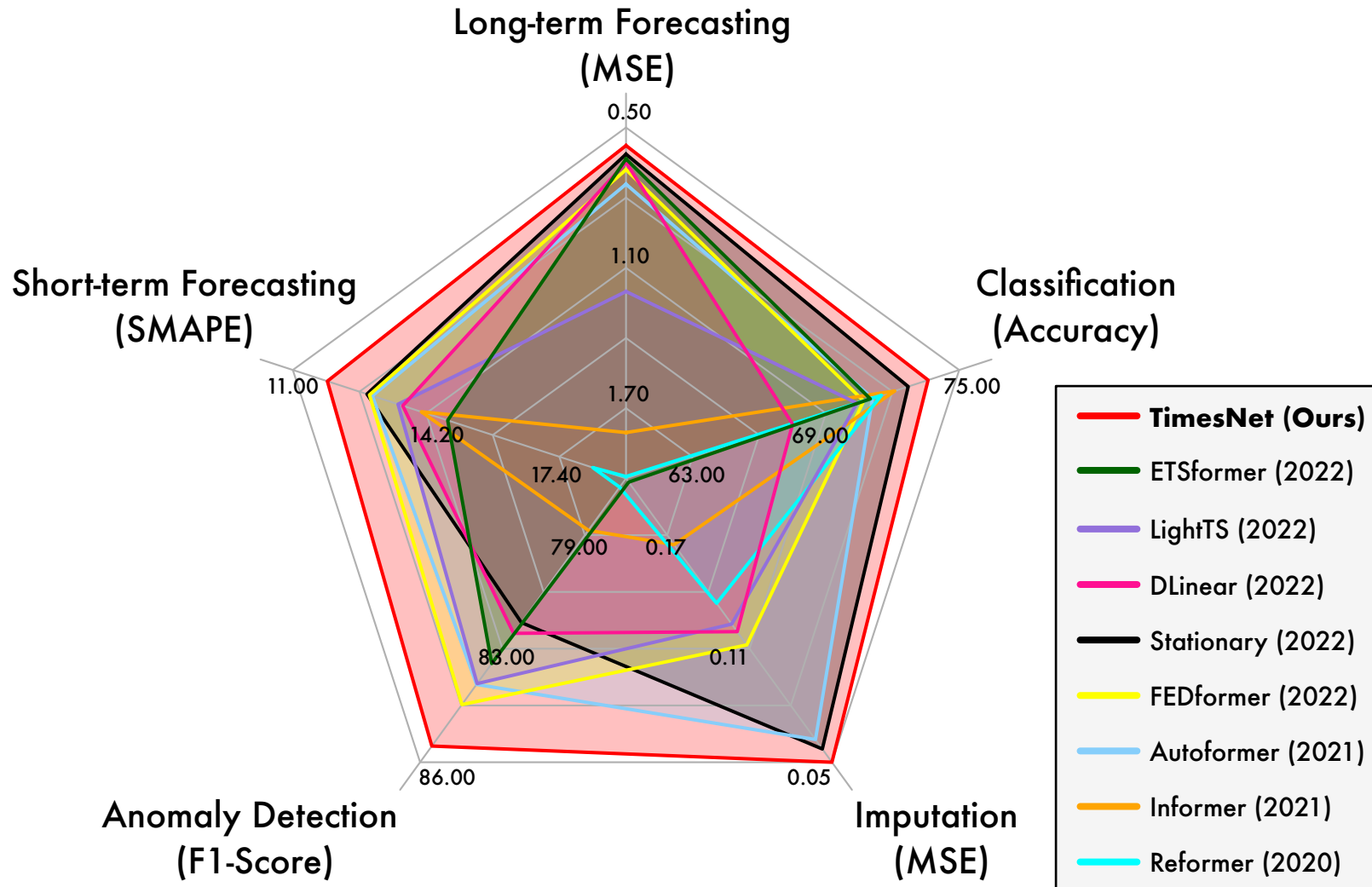
- ① 1D \rightarrow 2D
- ② 2D representation learning
- ③ 2D \rightarrow 1D

Experiment: Overall

Tasks	Benchmarks
Forecasting	Long-term: ETT (4 subsets), Electricity, Traffic, Weather, Exchange, ILI
	Short-term: M4 (6 subsets)
Imputation	ETT (4 subsets), Electricity, Weather
Classification	UEA (10 subsets)
Anomaly Detection	SMD, MSL, SMAP, SWaT, PSM

- ✓ Five mainstream time series analysis tasks.
- ✓ 36 datasets, 81 settings, 20+ baselines

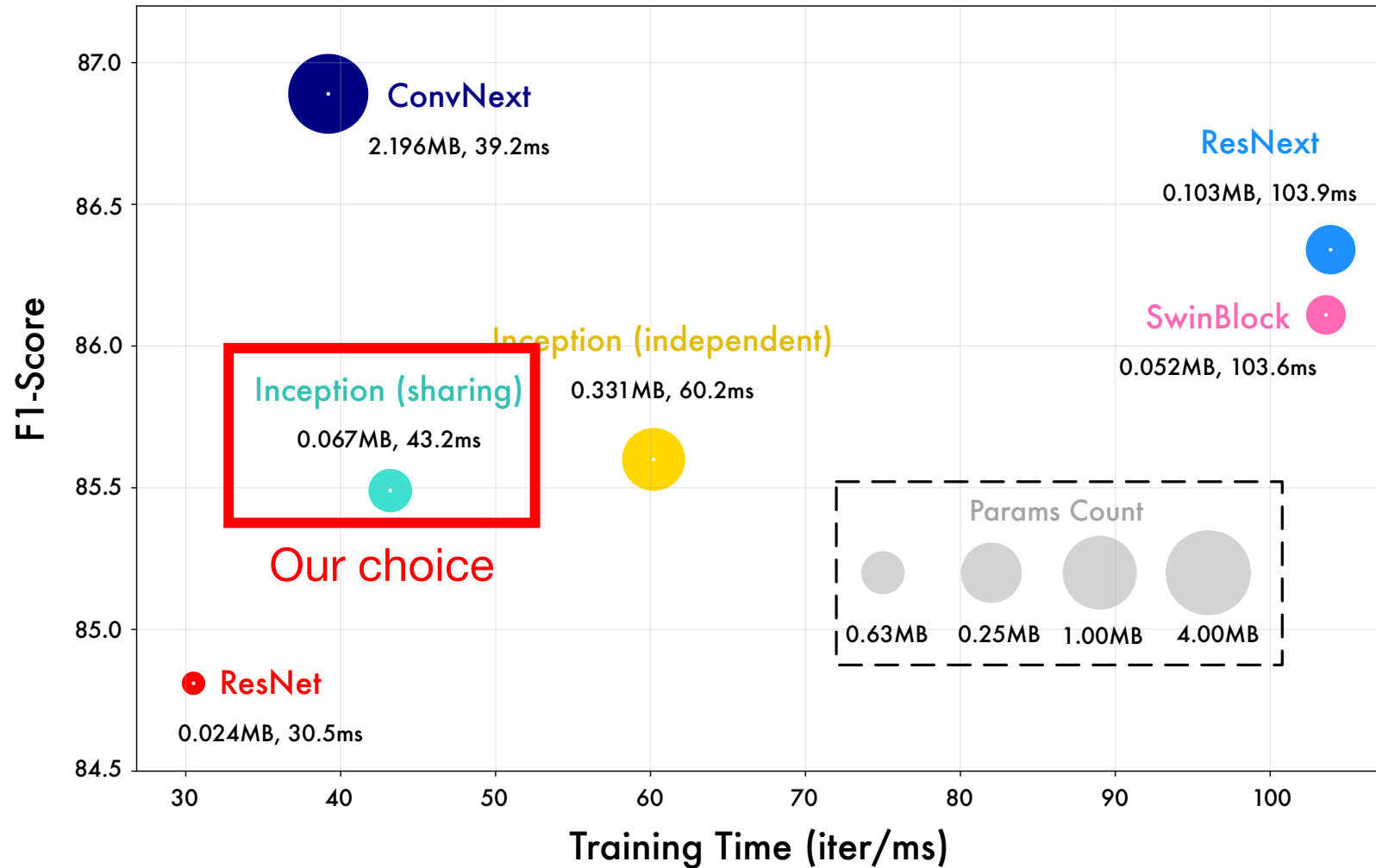
Experiment: Overall



TimesNet achieves state-of-the-art in all five tasks (2023/02)

Model Generality

Anomaly Detection



Better vision backbones,
Better performance 🏆

Bridge Time Series and
vision backbones 🏆

Time Series Library (TSlib)

The screenshot shows the GitHub repository page for 'Time Series Library (TSlib)' by user 'thuml'. The repository is public and has 3k stars, 514 forks, and 41 watchers. The main content area displays a list of files and folders with their commit history. The 'tutorial' folder is highlighted. The right sidebar contains sections for 'About', 'Releases', 'Packages', 'Contributors', and 'Languages'. The 'About' section describes it as a library for advanced deep time series models. The 'Contributors' section shows 22 contributors, with 8 highlighted. The 'Languages' section shows a bar chart for supported languages.

File/Folder	Description	Last Commit
data_provider	update tutorial about timeF encoding	3 months ago
exp	add TIDE and its ETTh1 script	3 weeks ago
layers	add n_heads for FEDformer	2 months ago
models	add imputation to TIDE	2 weeks ago
pic	update dataset discription	10 months ago
scripts	add imputation to TIDE	2 weeks ago
tutorial	docs: fix and improve content in tutorial/TimesNet_tutoria...	2 months ago
utils	capitalization	2 weeks ago
.gitignore	Official implementation of iTransformer added	2 months ago
LICENSE	init	10 months ago
README.md	Update README.md	4 days ago
requirements.txt	clean	9 months ago
run.py	Merge pull request #276 from jurij-ch/argsprint	2 weeks ago

About
A Library for Advanced Deep Time Series Models.
deep-learning time-series
Readme
MIT license
Activity
3k stars
41 watching
514 forks
Report repository

Releases
No releases published
[Create a new release](#)

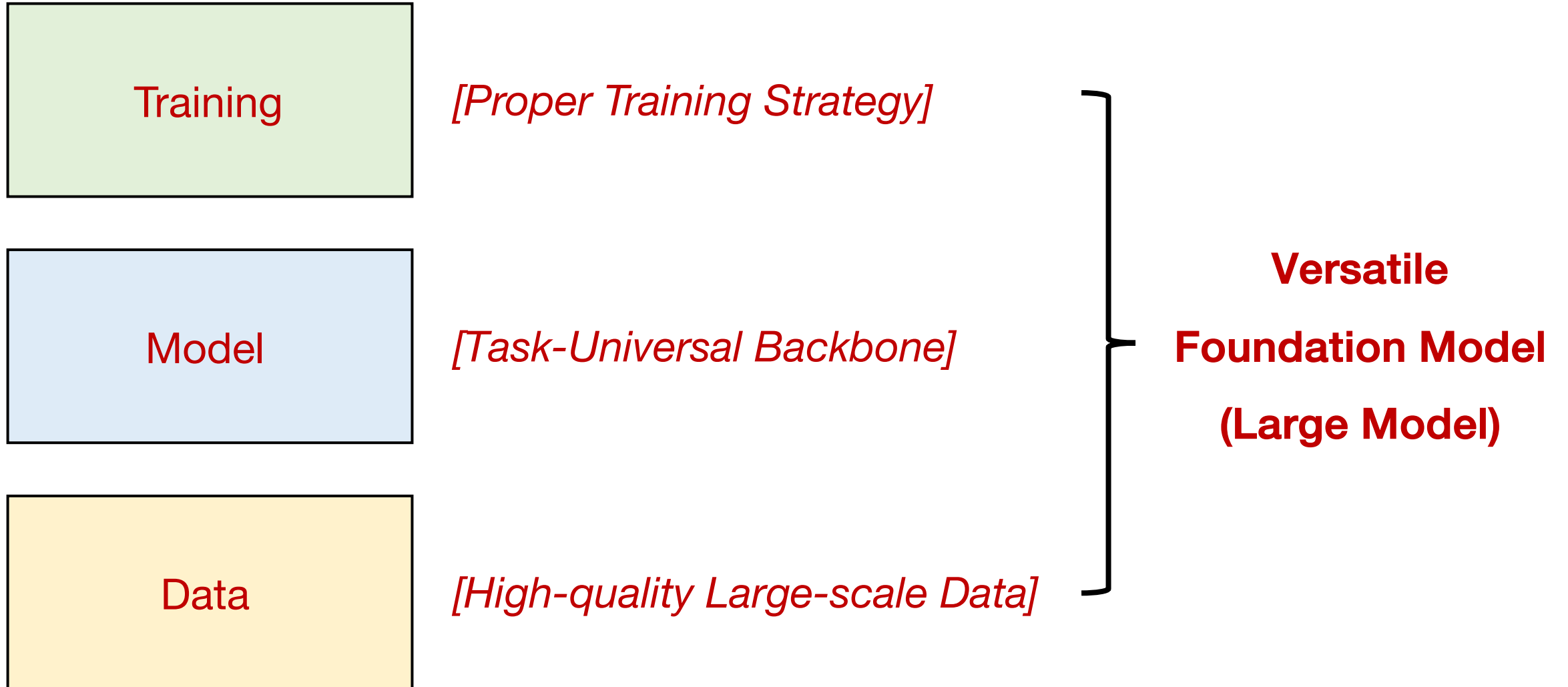
Packages
No packages published
[Publish your first package](#)

Contributors 22
+ 8 contributors

Languages

Code is available at <https://github.com/thuml/Time-Series-Library> with 3000+ stars

Foundation Models for Time Series





2023 CCF国际AIOps挑战赛决赛暨“大模型时代的AIOps”研讨会

THANKS