



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



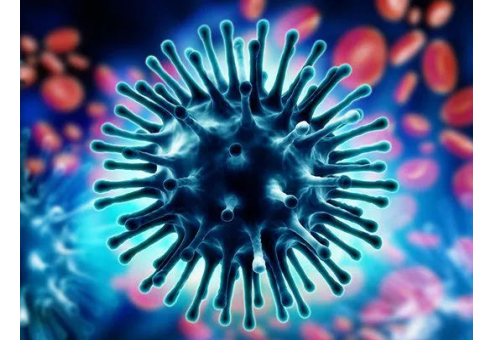
Traffic  
Flow



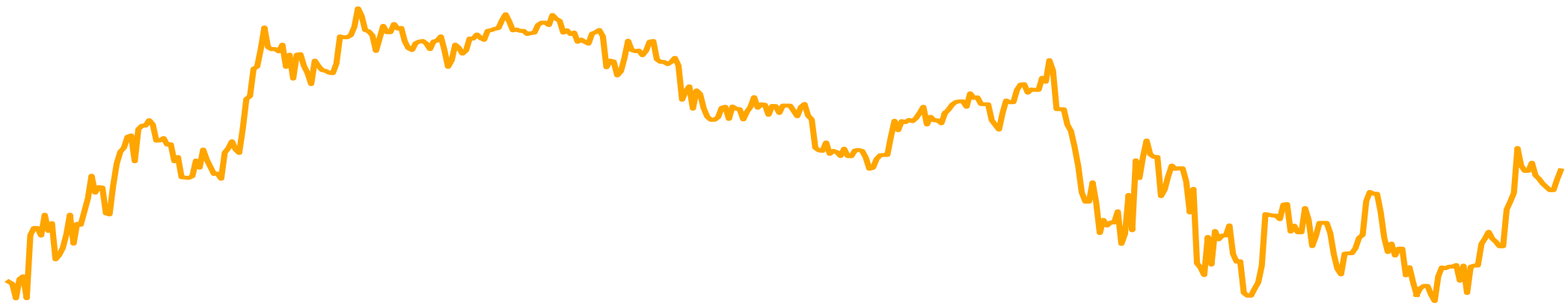
Economic  
Changes



Weather  
Variations

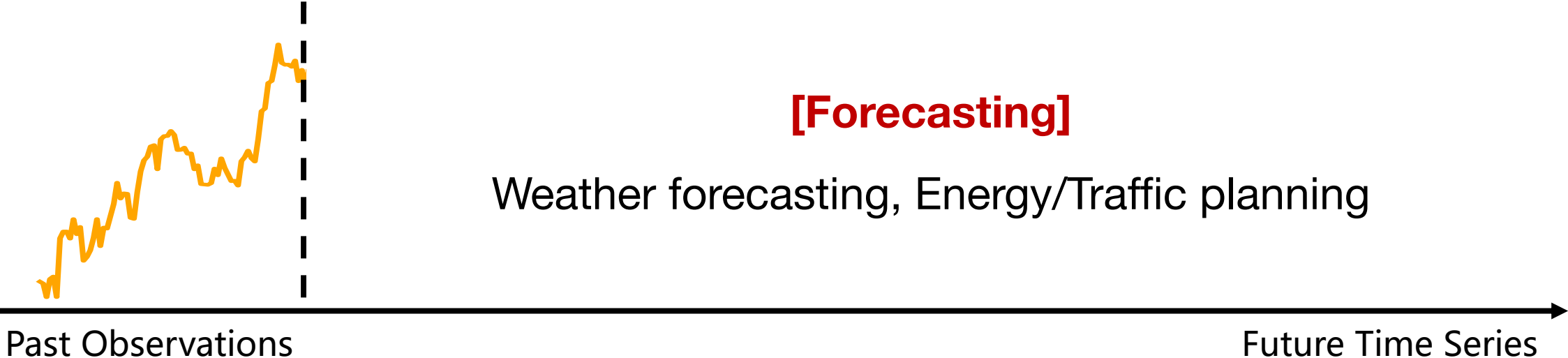


Disease  
Propagation

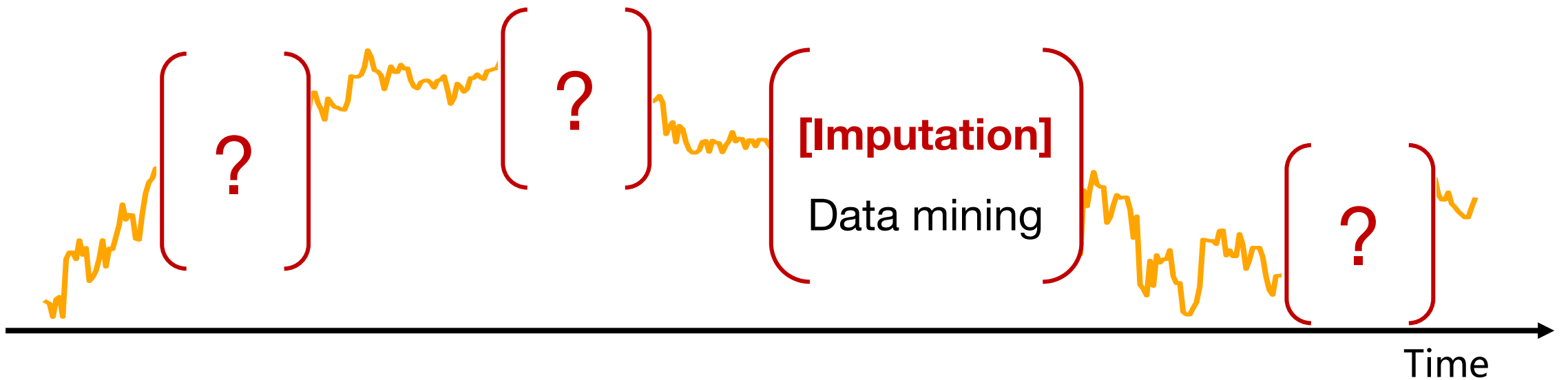
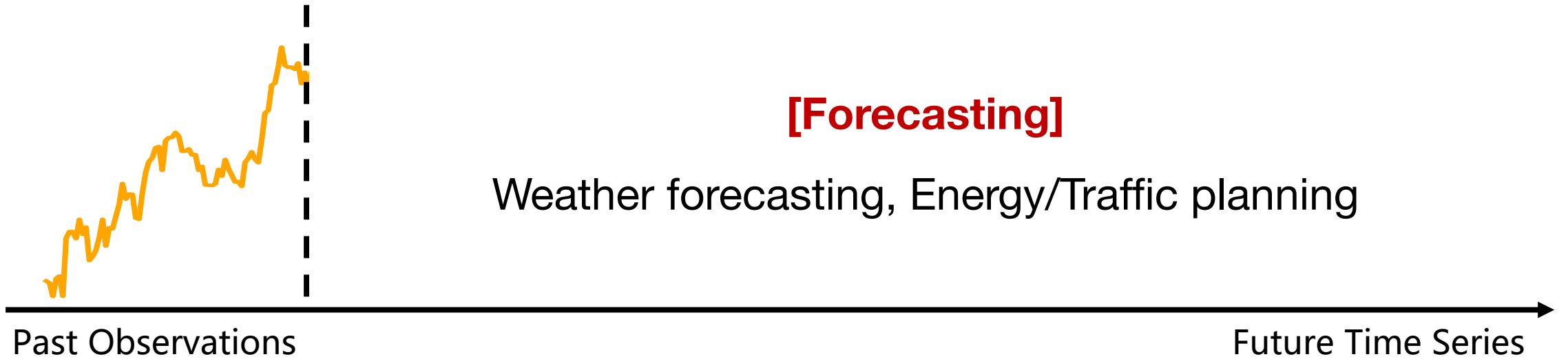


Time

# Time Series Analysis



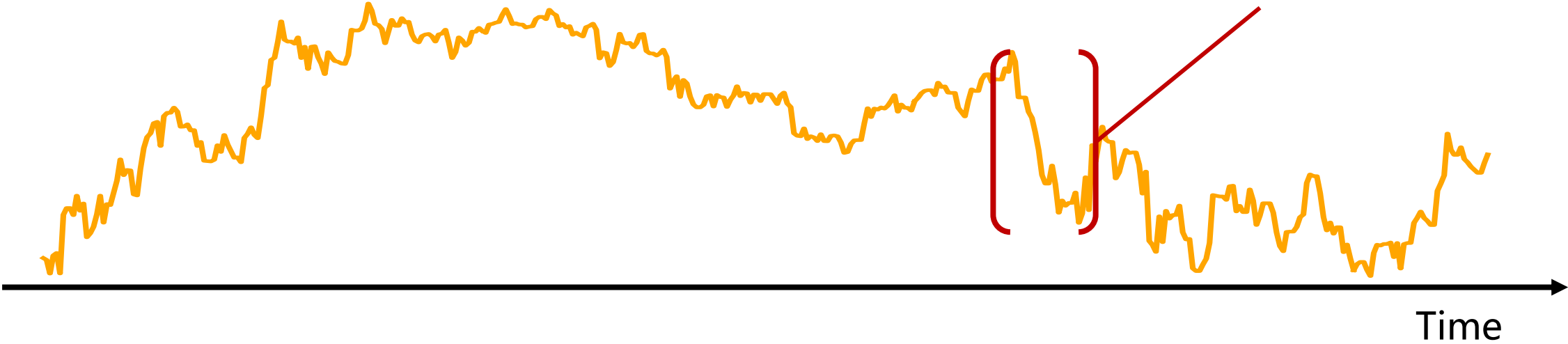
# Time Series Analysis



# Time Series Analysis

**[Anomaly Detection]**

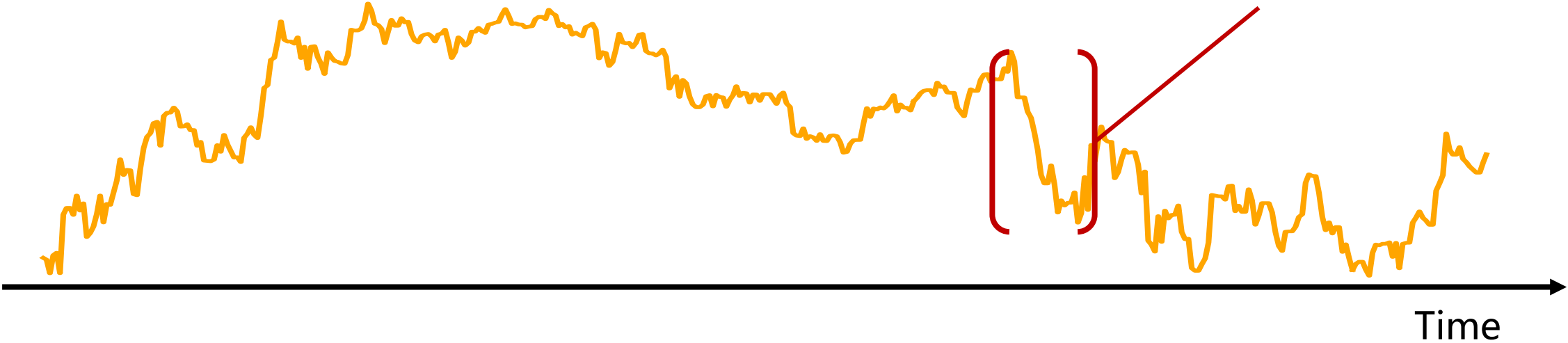
Industrial Maintenance



# Time Series Analysis

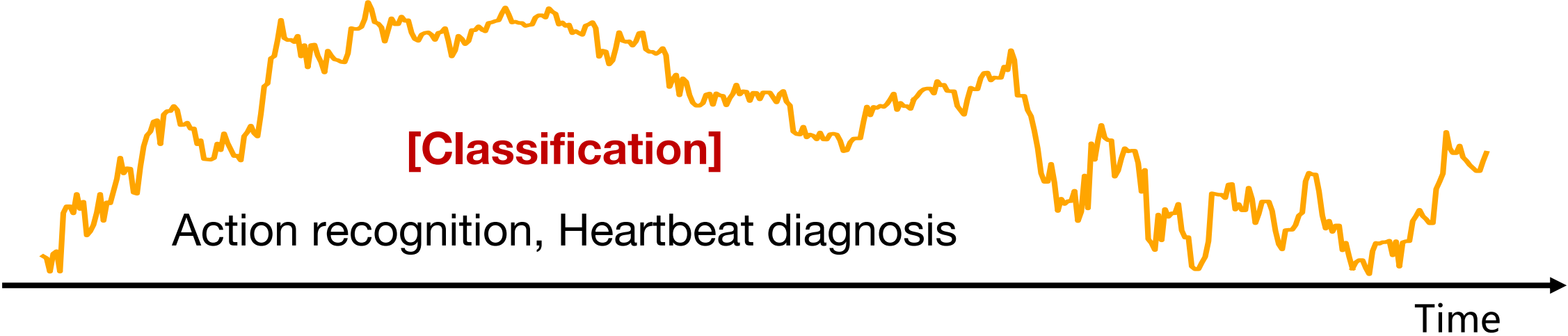
**[Anomaly Detection]**

Industrial Maintenance

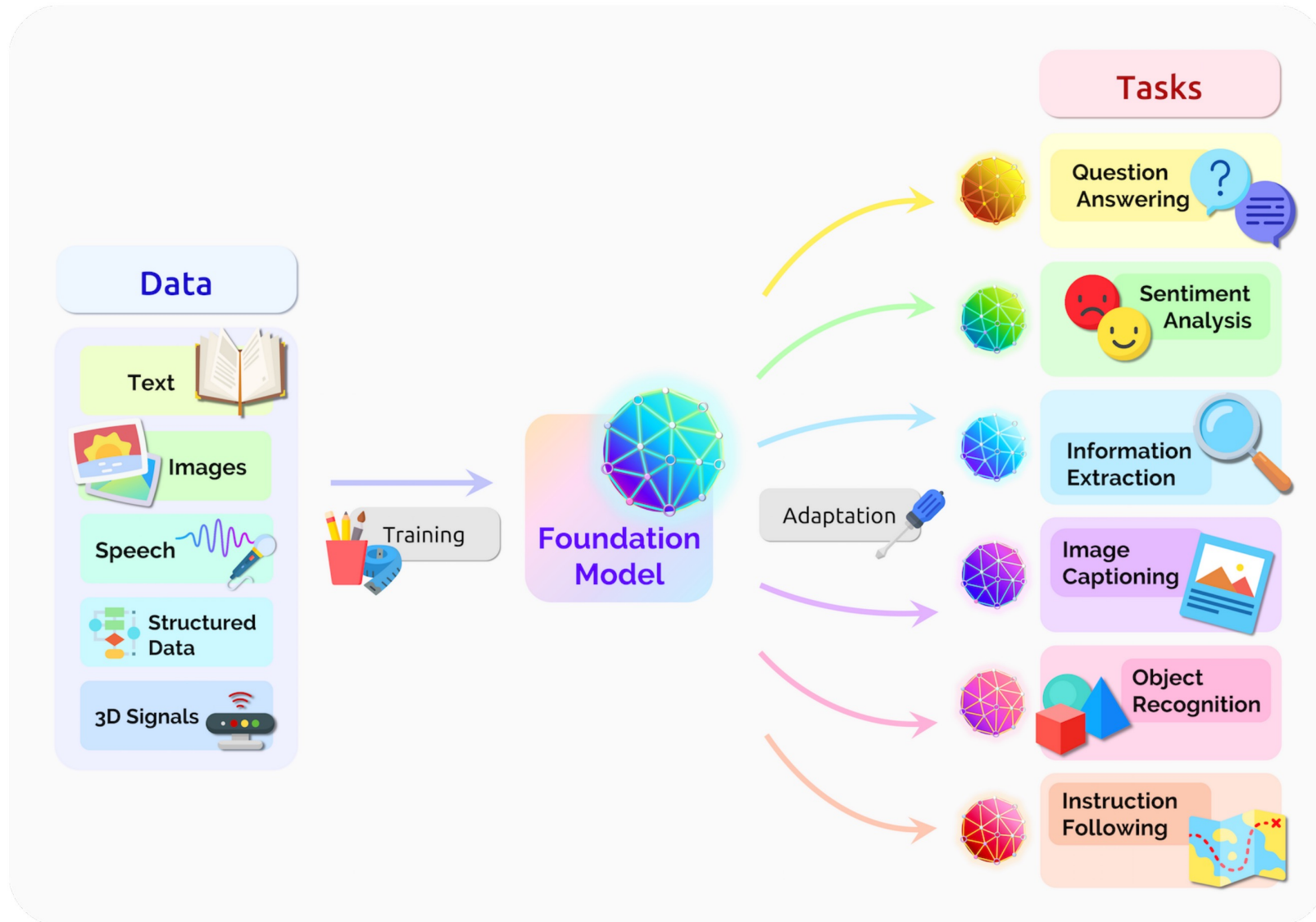


**[Classification]**

Action recognition, Heartbeat diagnosis



# In Pursuing Foundation Models



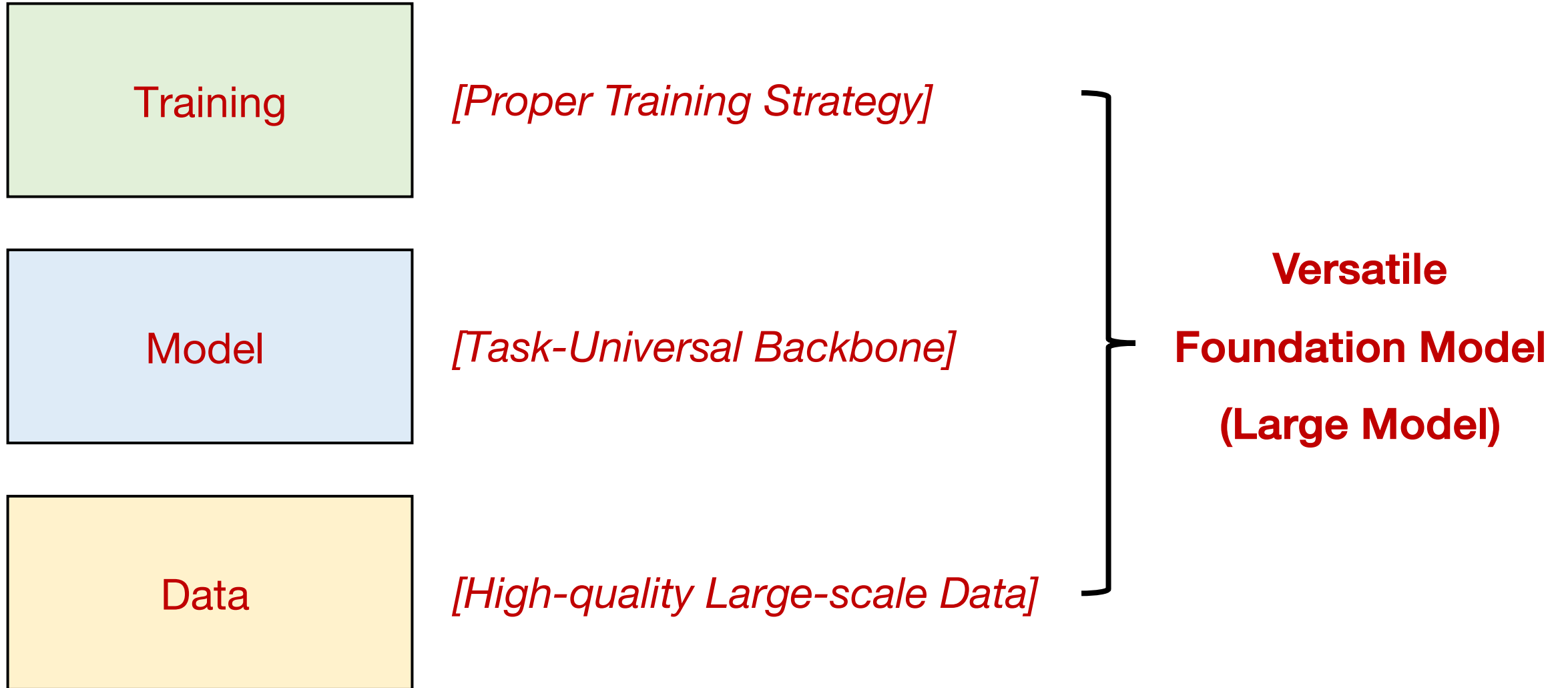
***[Data Universal]***

Learn from various modalities

***[Task Universal]***

Adapt to a wide range of  
downstream tasks

# In Pursuing Foundation Models







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



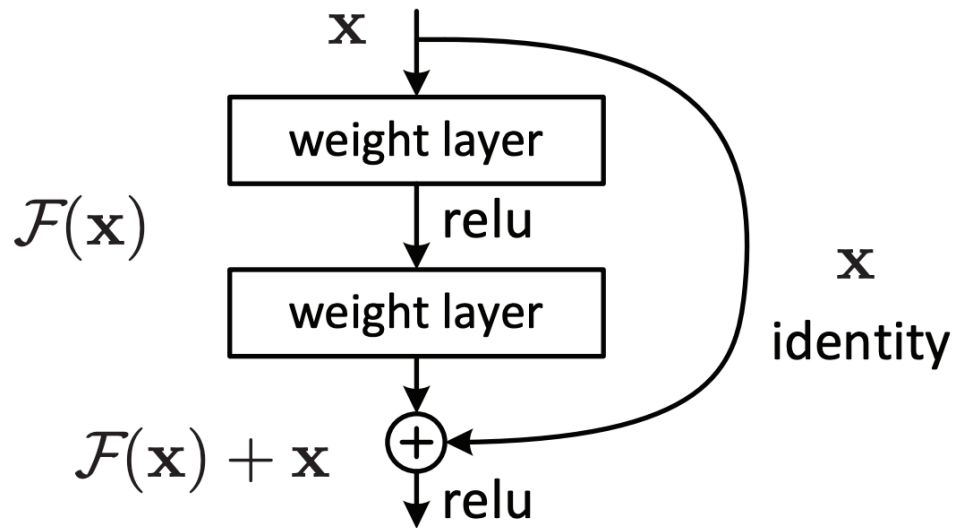
Jianmin Wang



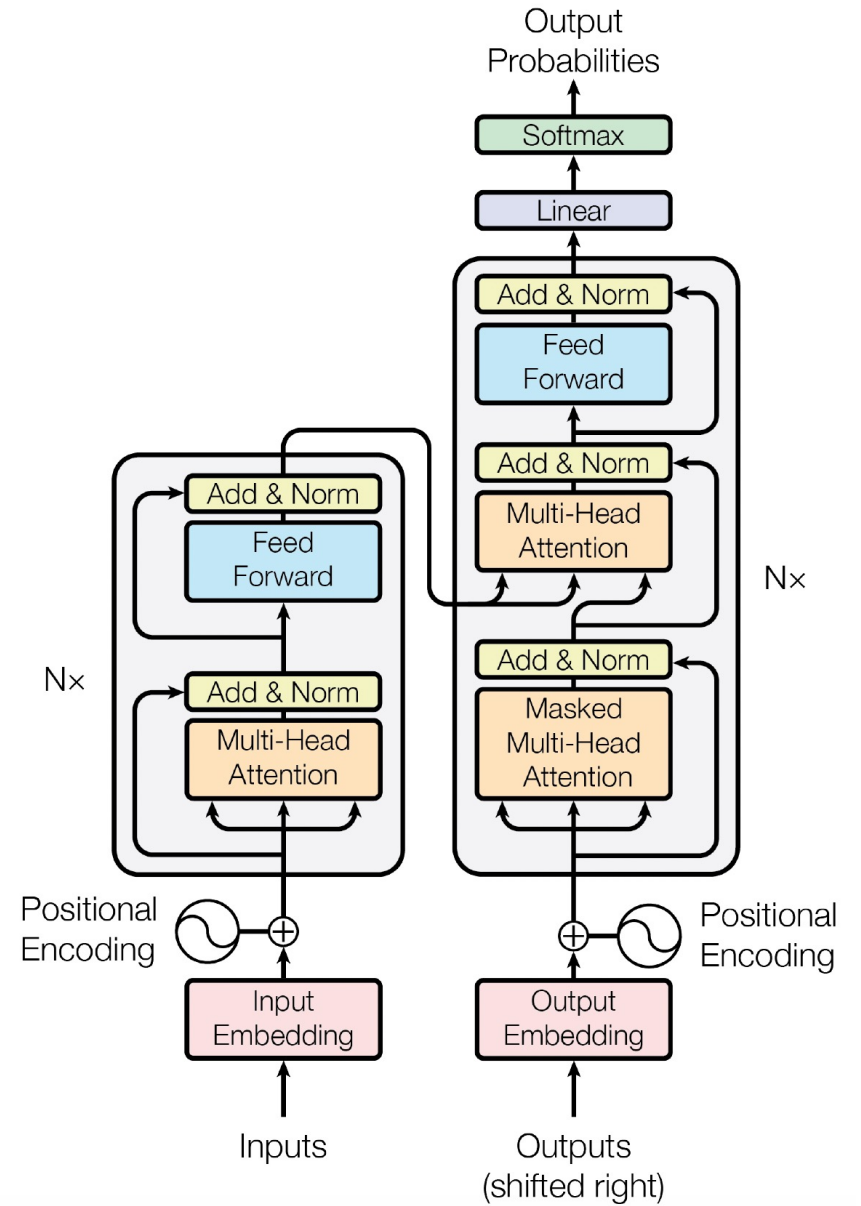
Mingsheng Long

# Foundation Models in CV and NLP

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



Classification, Object detection, Segmentation

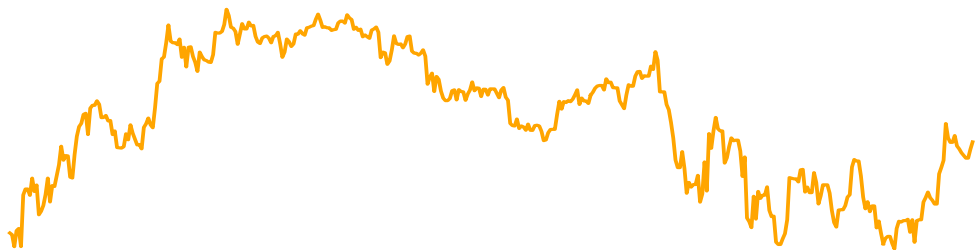


Classification, Generation

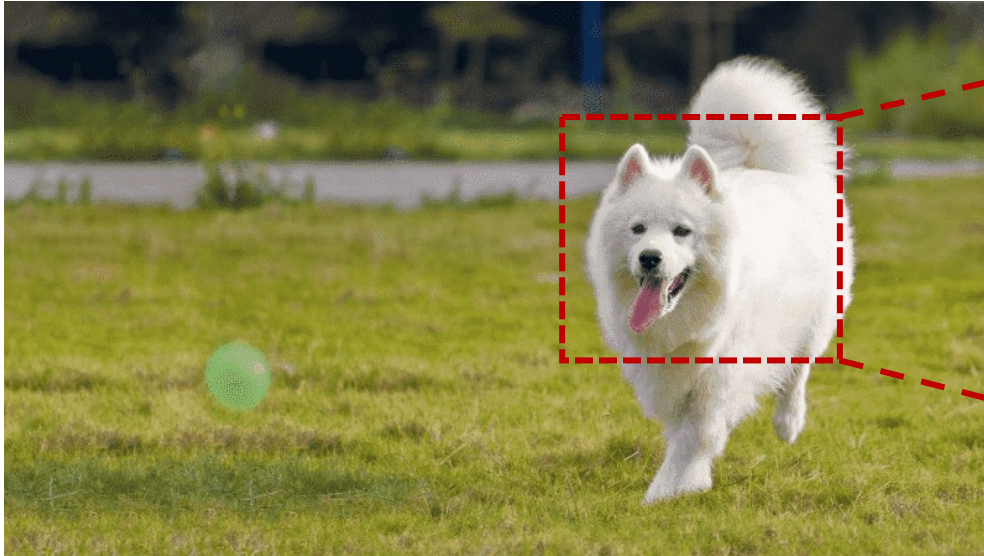
# Differences among Image, Language, Time Series



TimesNet is for time series analysis.

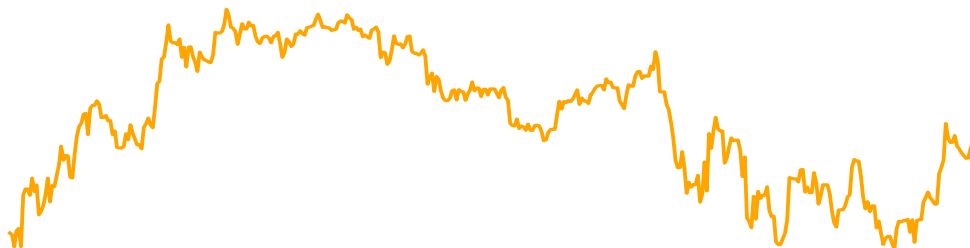


# Differences among Image, Language, Time Series



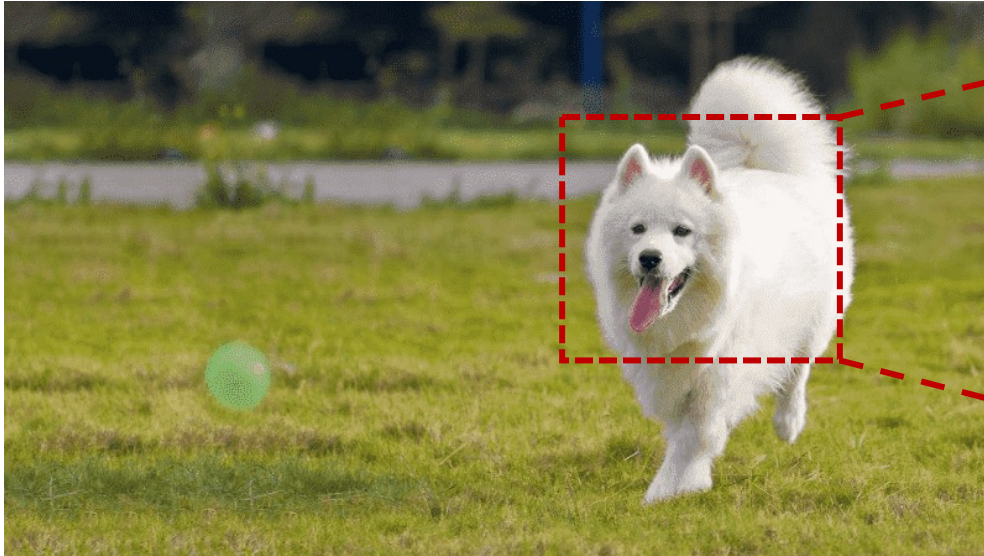
TimesNet is for time series analysis.

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



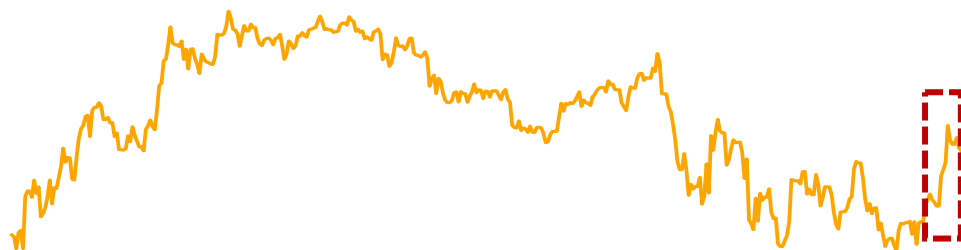
WIKIPEDIA  
The Free Encyclopedia

# 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.



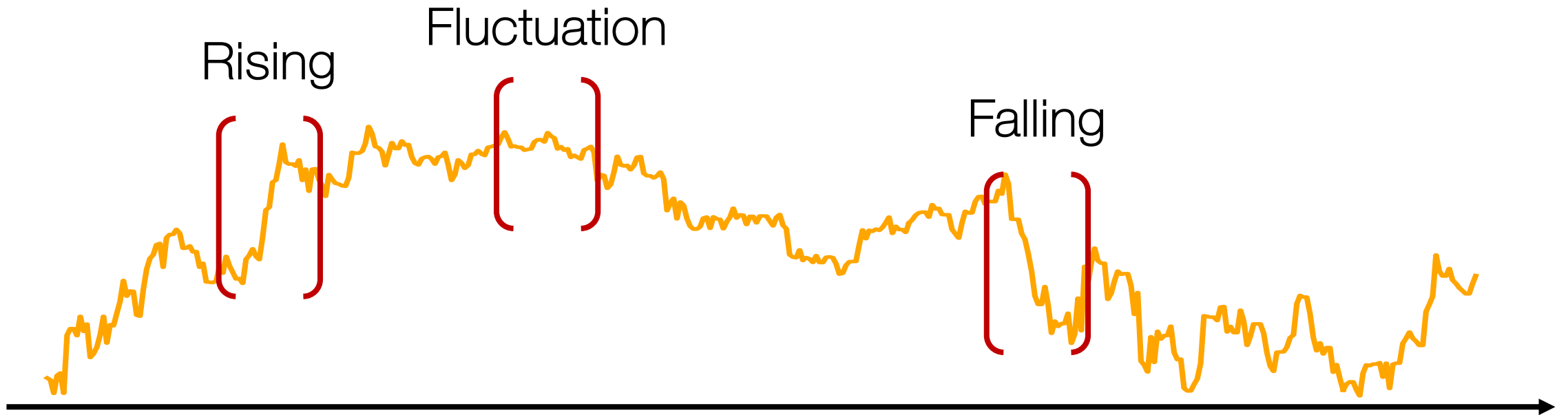
**Each time point only saves some scalars.**



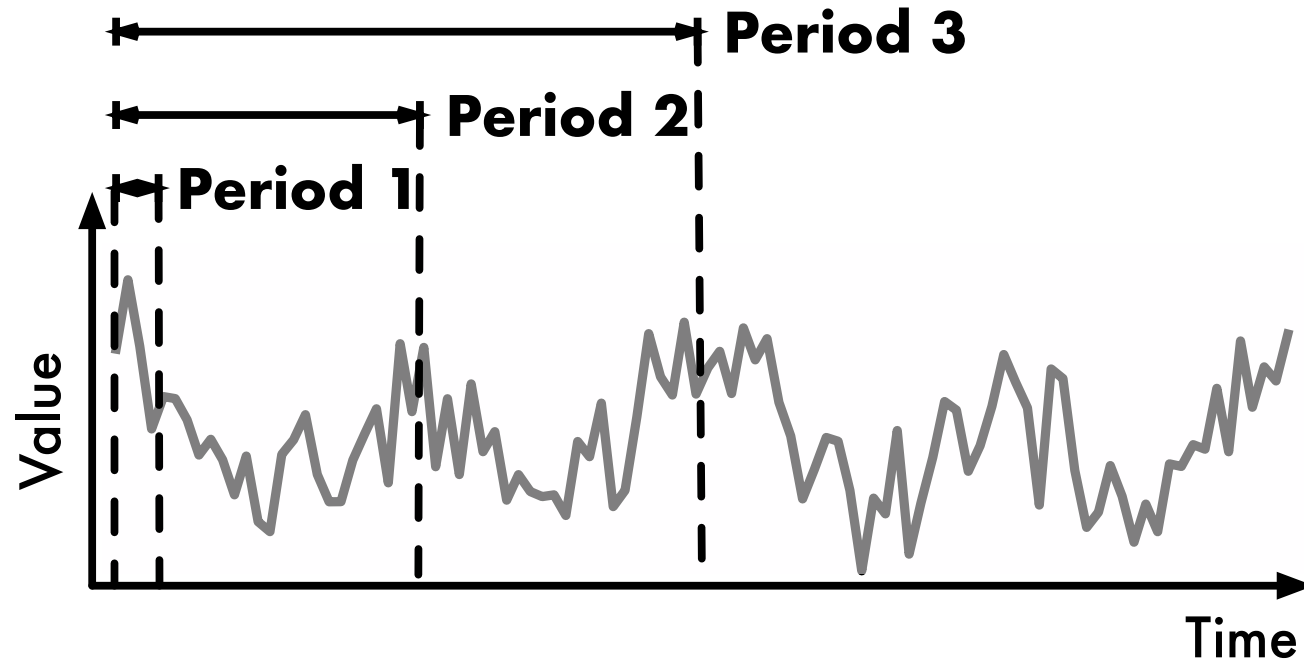
WIKIPEDIA  
The Free Encyclopedia

# 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

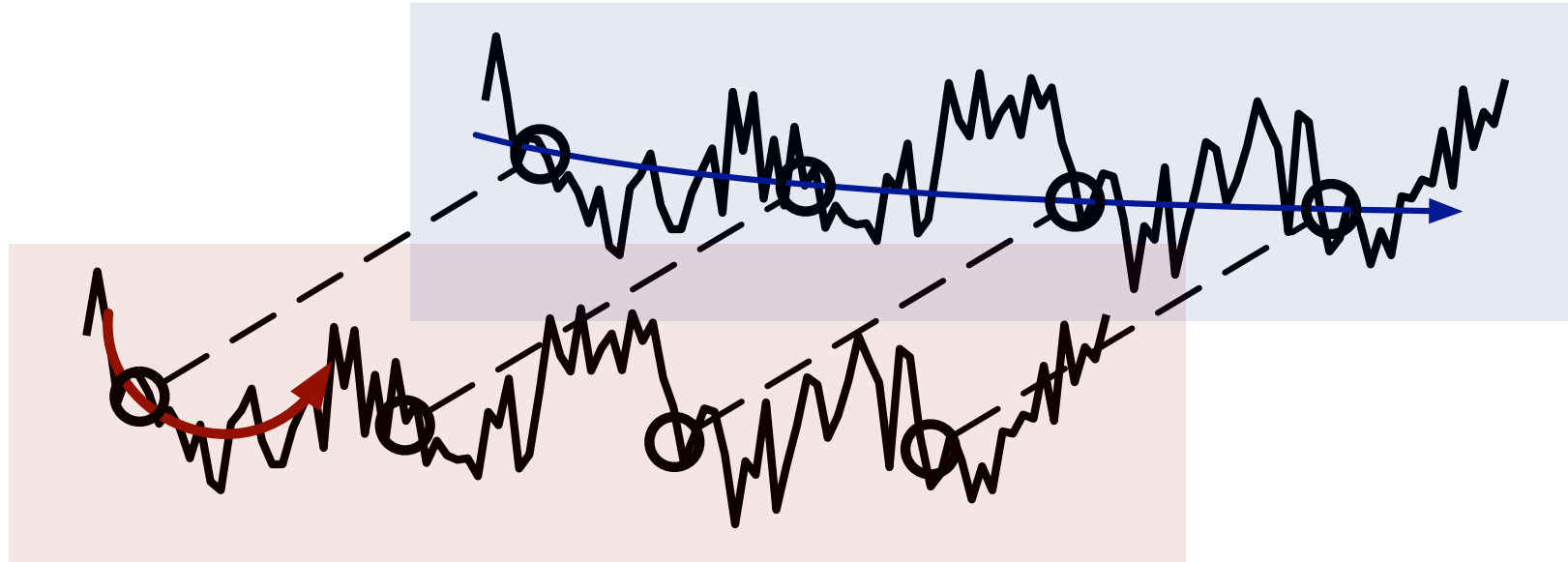


- ✓ Traffic: daily and weekly
- ✓ Weather: daily and yearly

**Real-world time series usually present multi-periodicity.**

**Multiple periods overlap and interact with each other.**

# Intraperiod- and Interperiod-variations

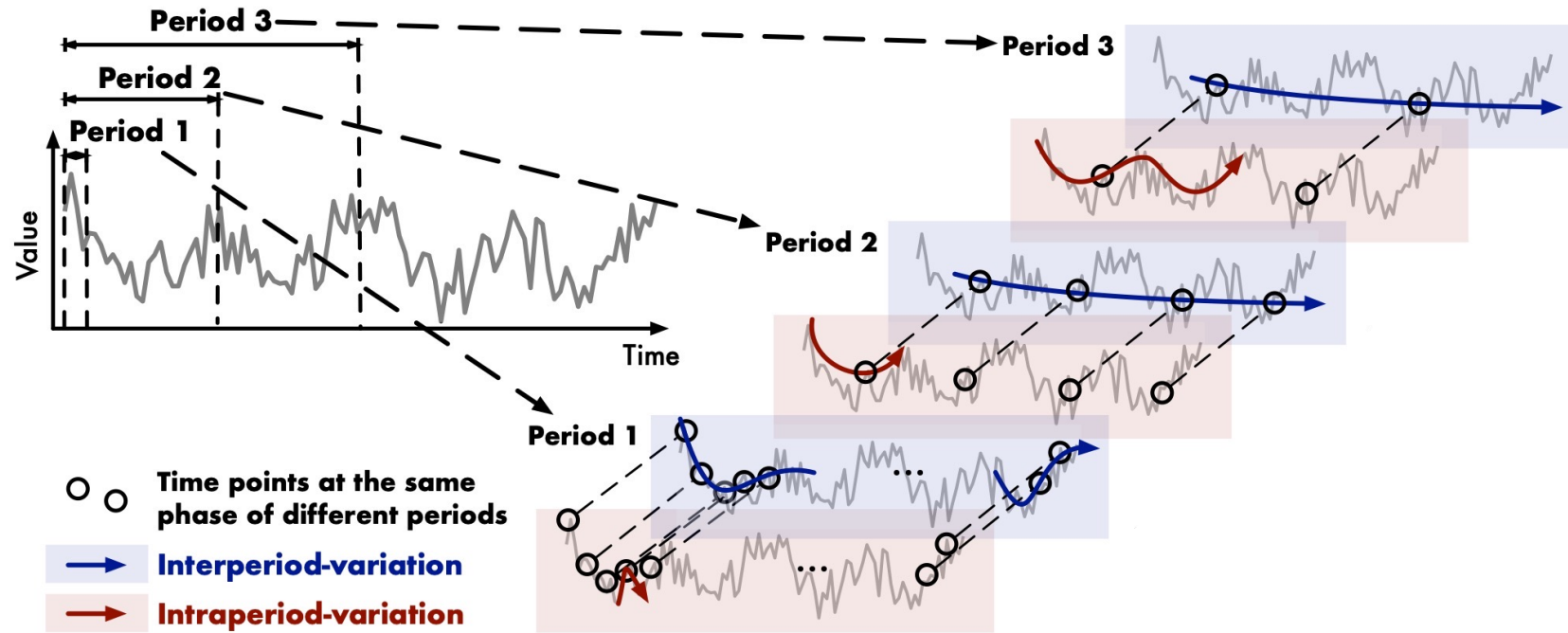


- ✓ **Intraperiod**: adjacent area, **short-term variations**
- ✓ **Interperiod**: same phase in adjacent periods, **long-term variations**

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



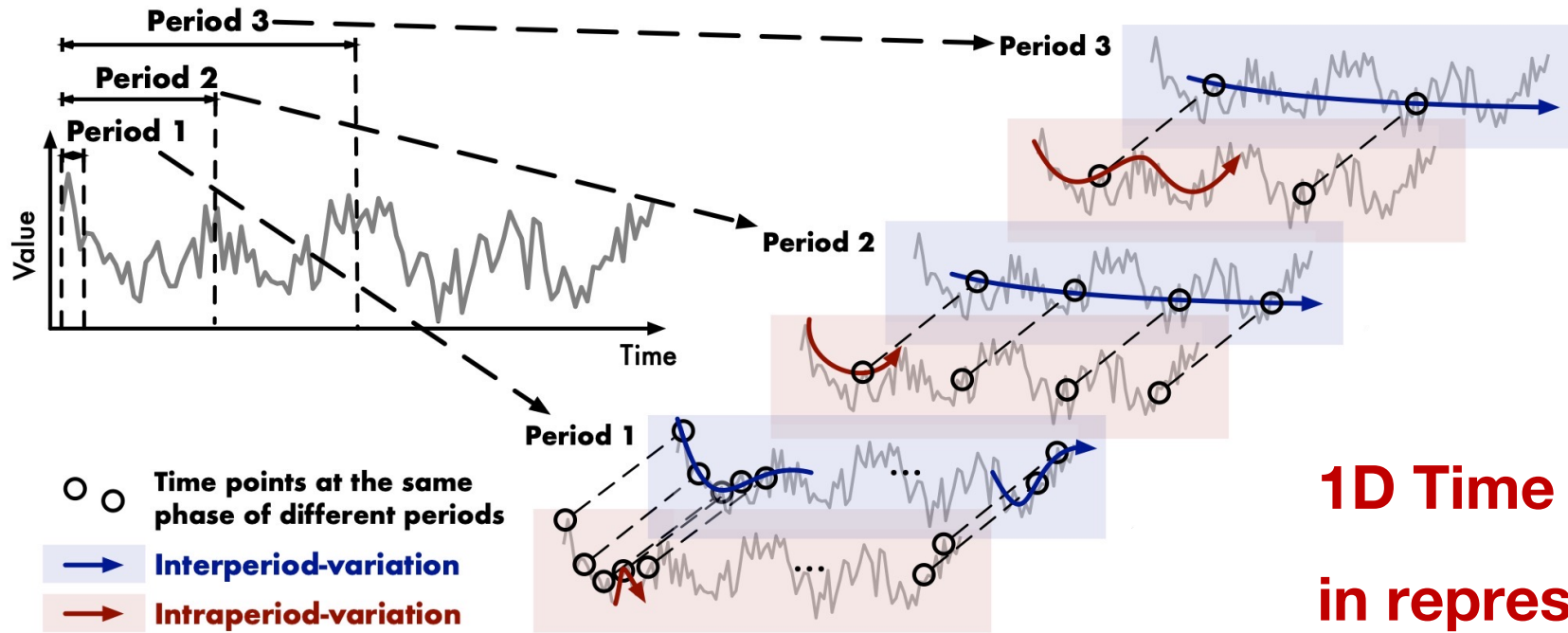
# Overall design of TimesNet



## ① Multi-periodicity

A modular architecture to disentangle intricate temporal patterns

# Overall design of TimesNet

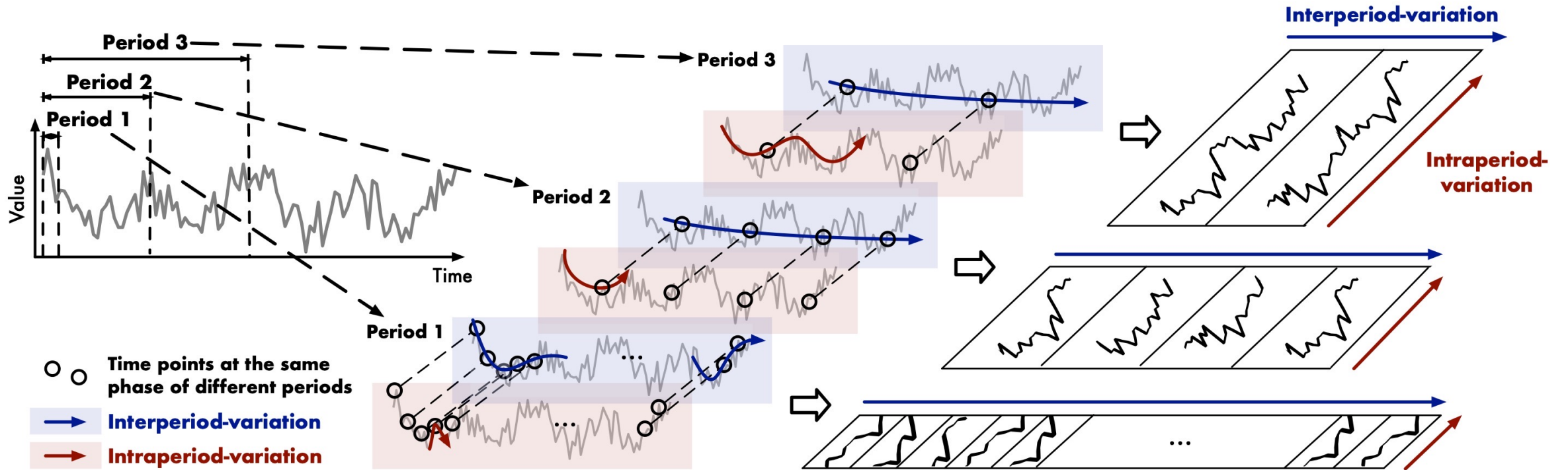


**1D Time Series has limitations in representation capability.**

## ① Multi-periodicity

A modular architecture to disentangle intricate temporal patterns

# Overall design of TimesNet

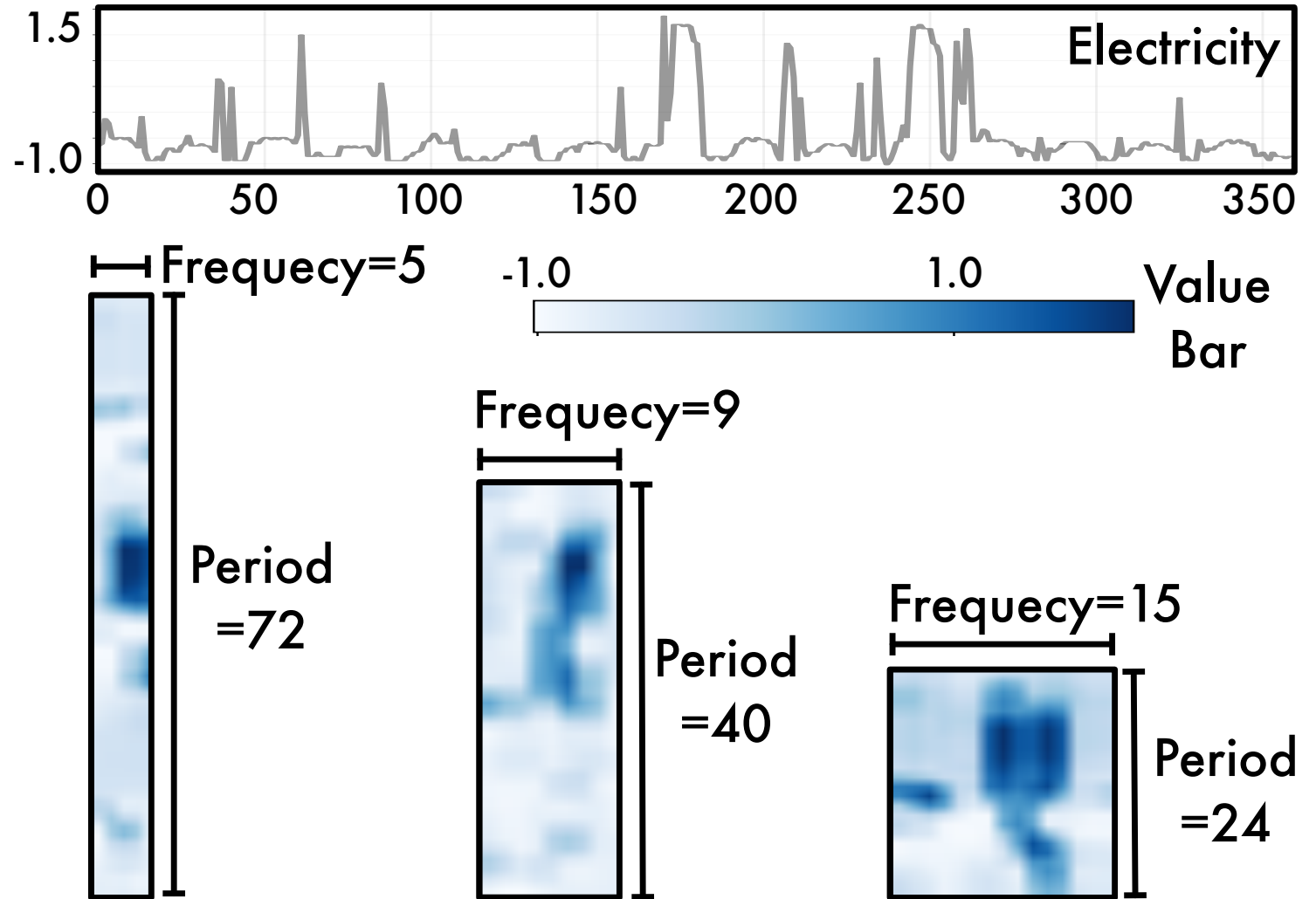


① Multi-periodicity    ② Temporal 2D-variation

Unify intraperiod- and interperiod-variations in 2D space by **reshape**

# Temporal 2D-variation: A Case Study

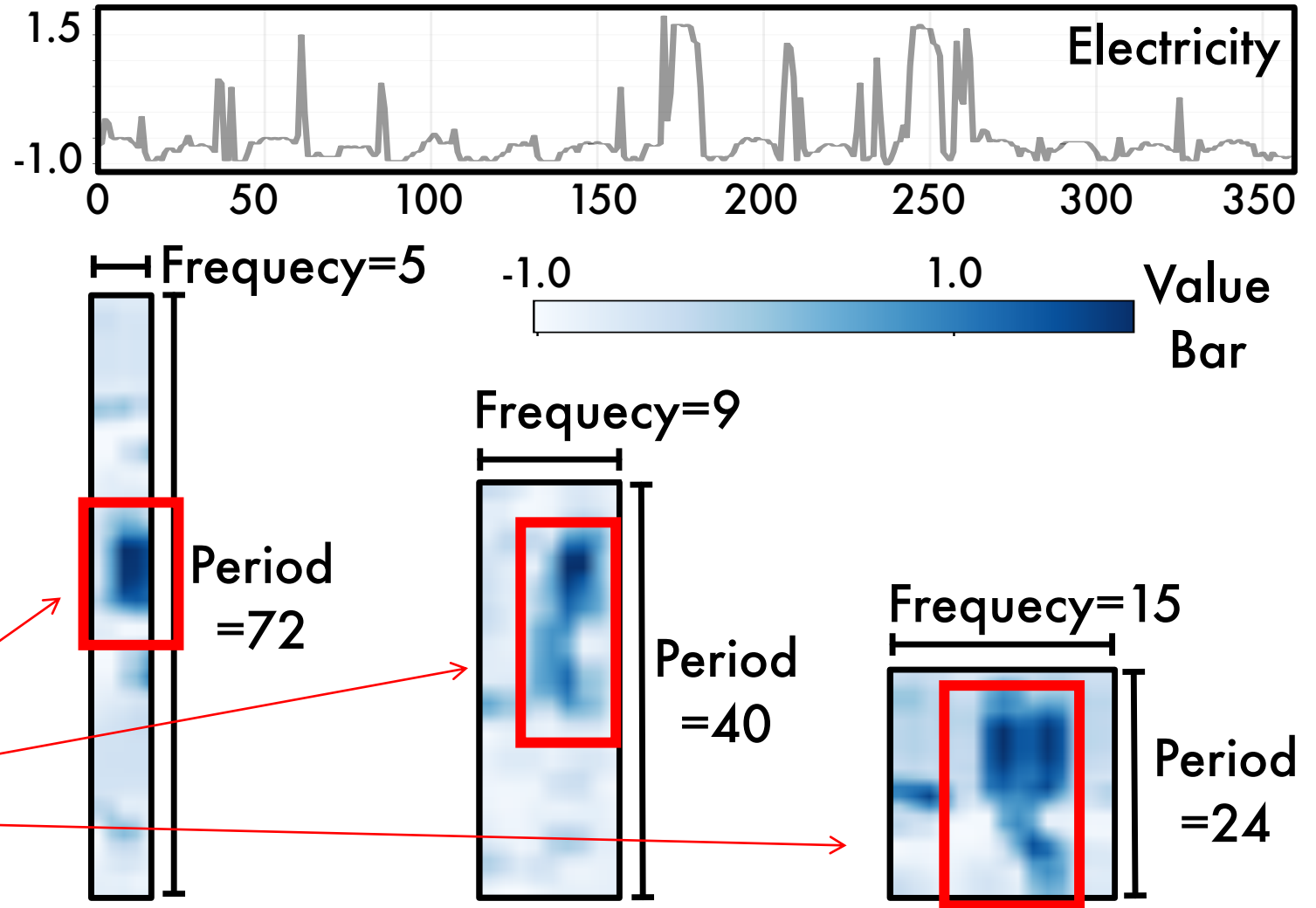
- ✓ Reshape the 1D time series into 2D according to periods.
- ✓ Two dimensions represent interperiod- and intraperiod-variations respectively.



# Temporal 2D-variation: A Case Study

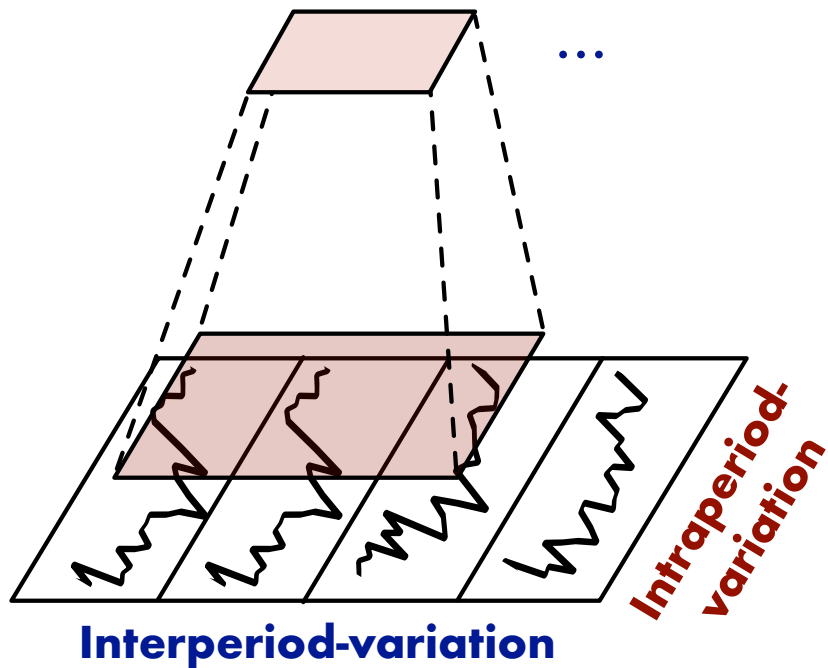
- ✓ Reshape the 1D time series into 2D according to periods.
- ✓ Two dimensions represent interperiod- and intraperiod-variations respectively.

2D locality



# Temporal 2D-variation: A Case Study

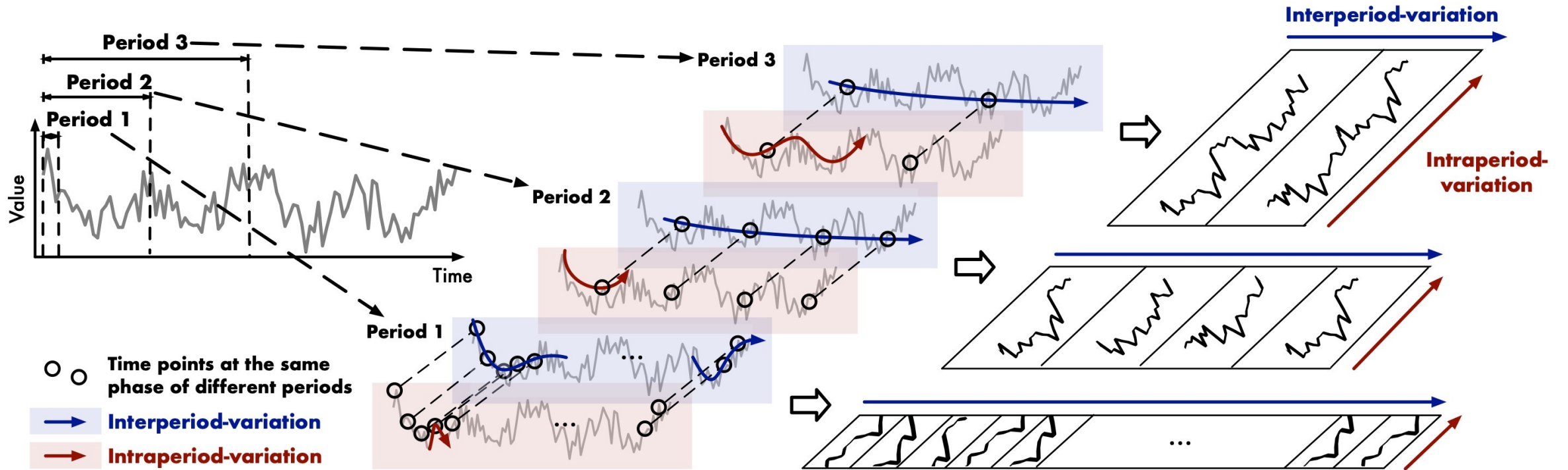
Capture Temporal 2D-variations  
by 2D Kernels



With temporal 2D-variations, we can

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

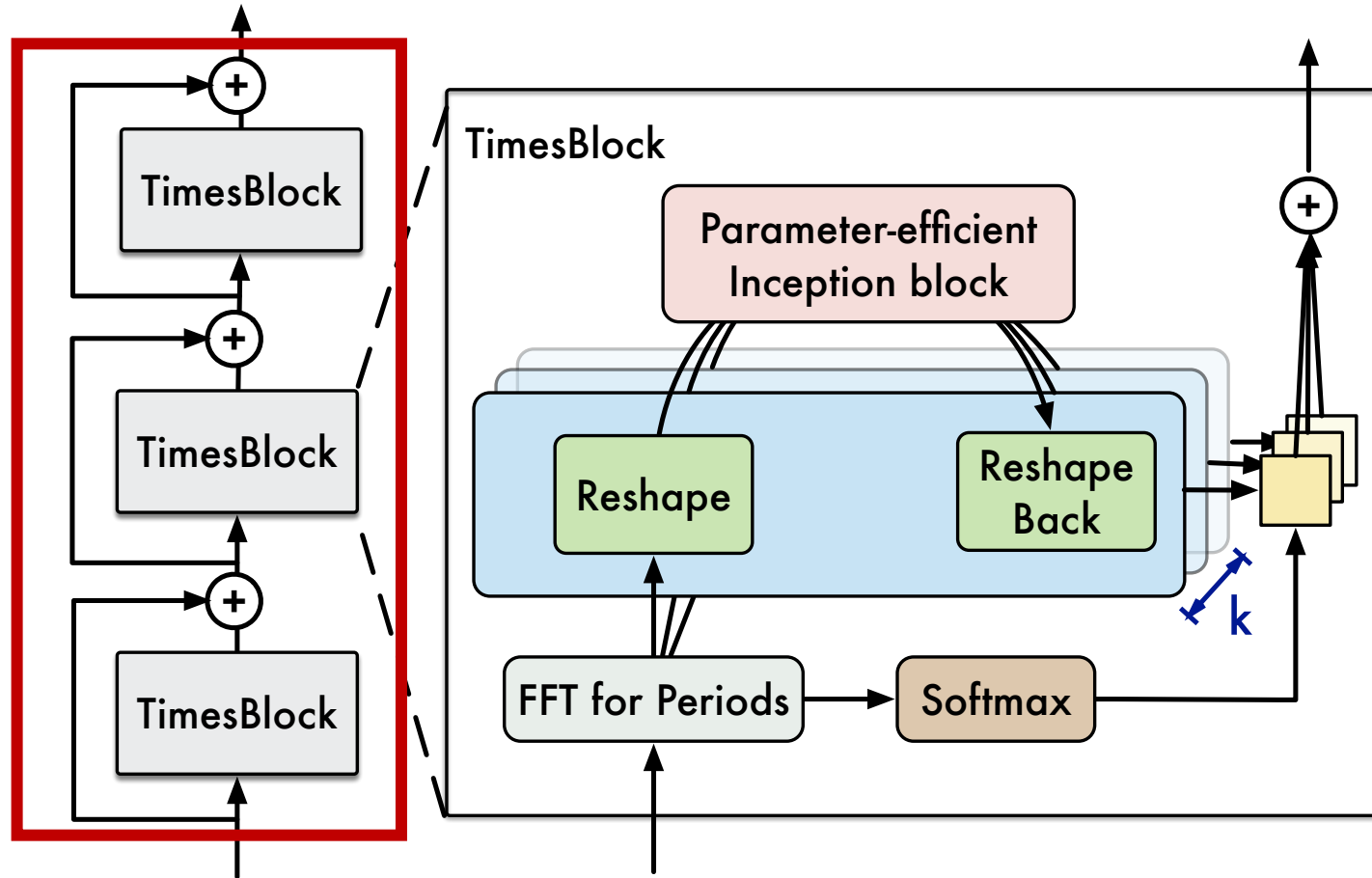
# Overall design of TimesNet



① Multi-periodicity    ② Temporal 2D-variation

Unify intraperiod- and interperiod-variations in 2D

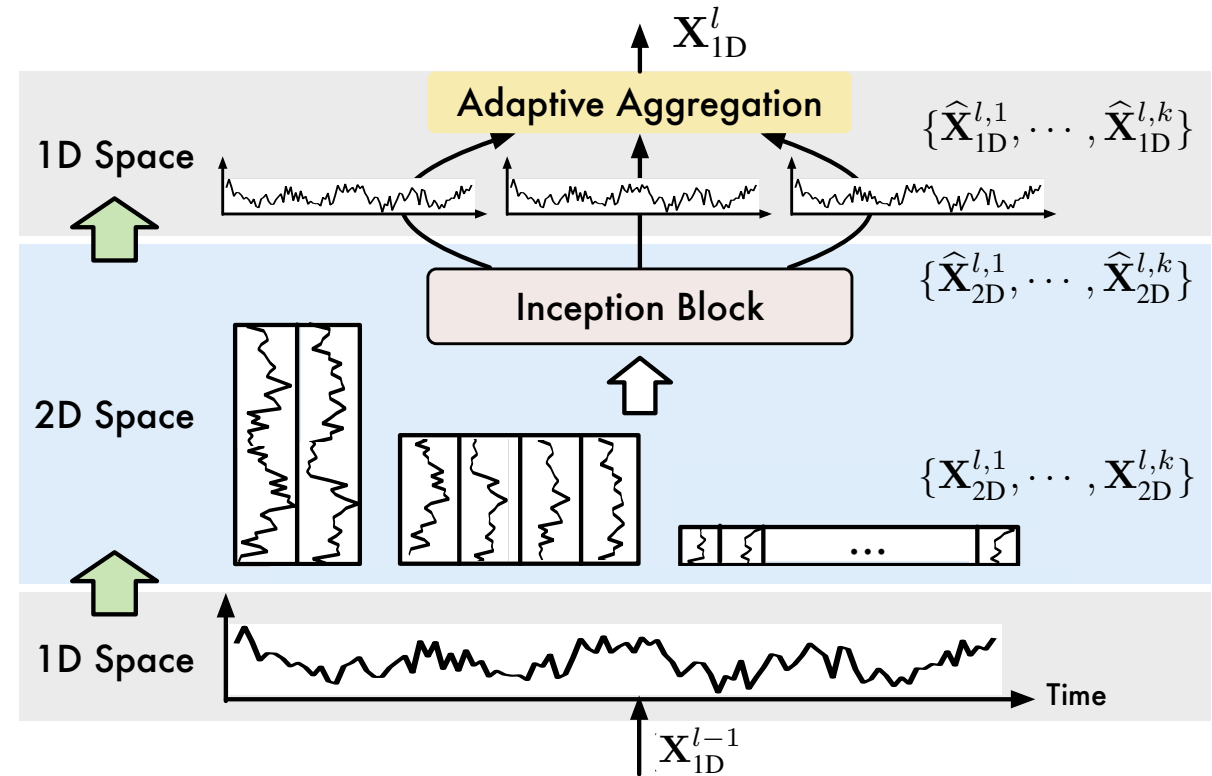
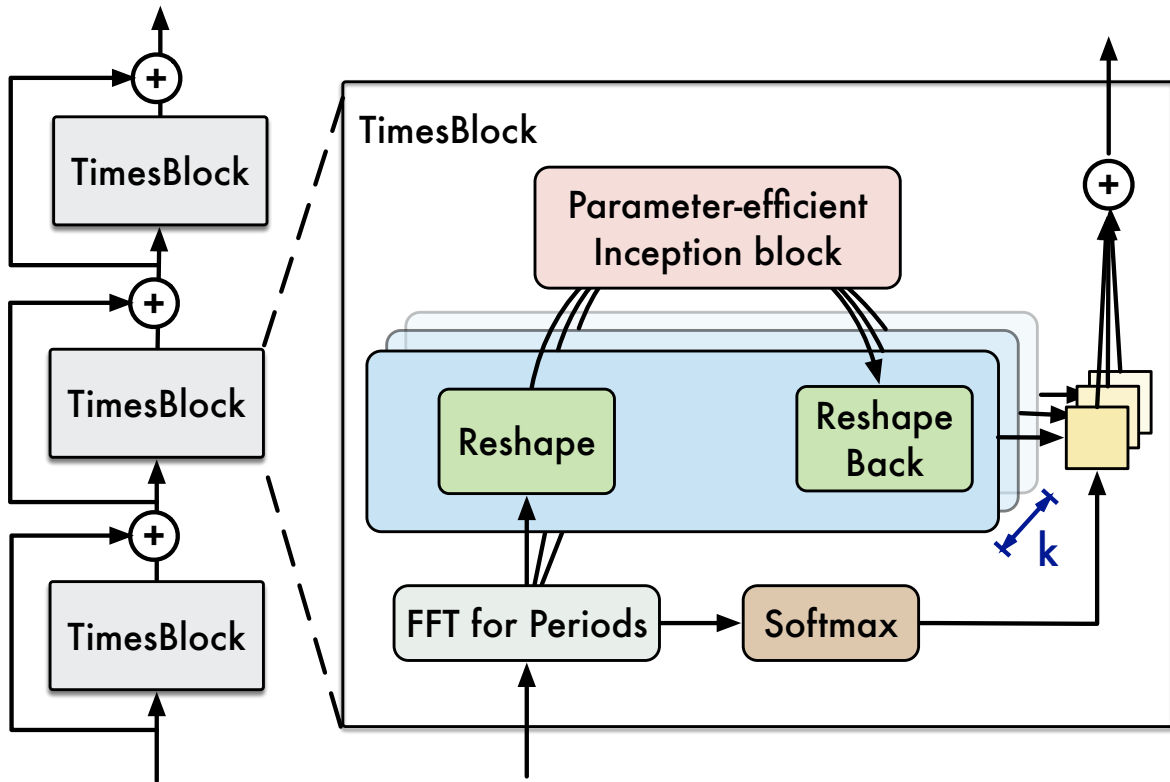
# TimesNet



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



# TimesBlock



TimesBlock learns representations in 2D space.

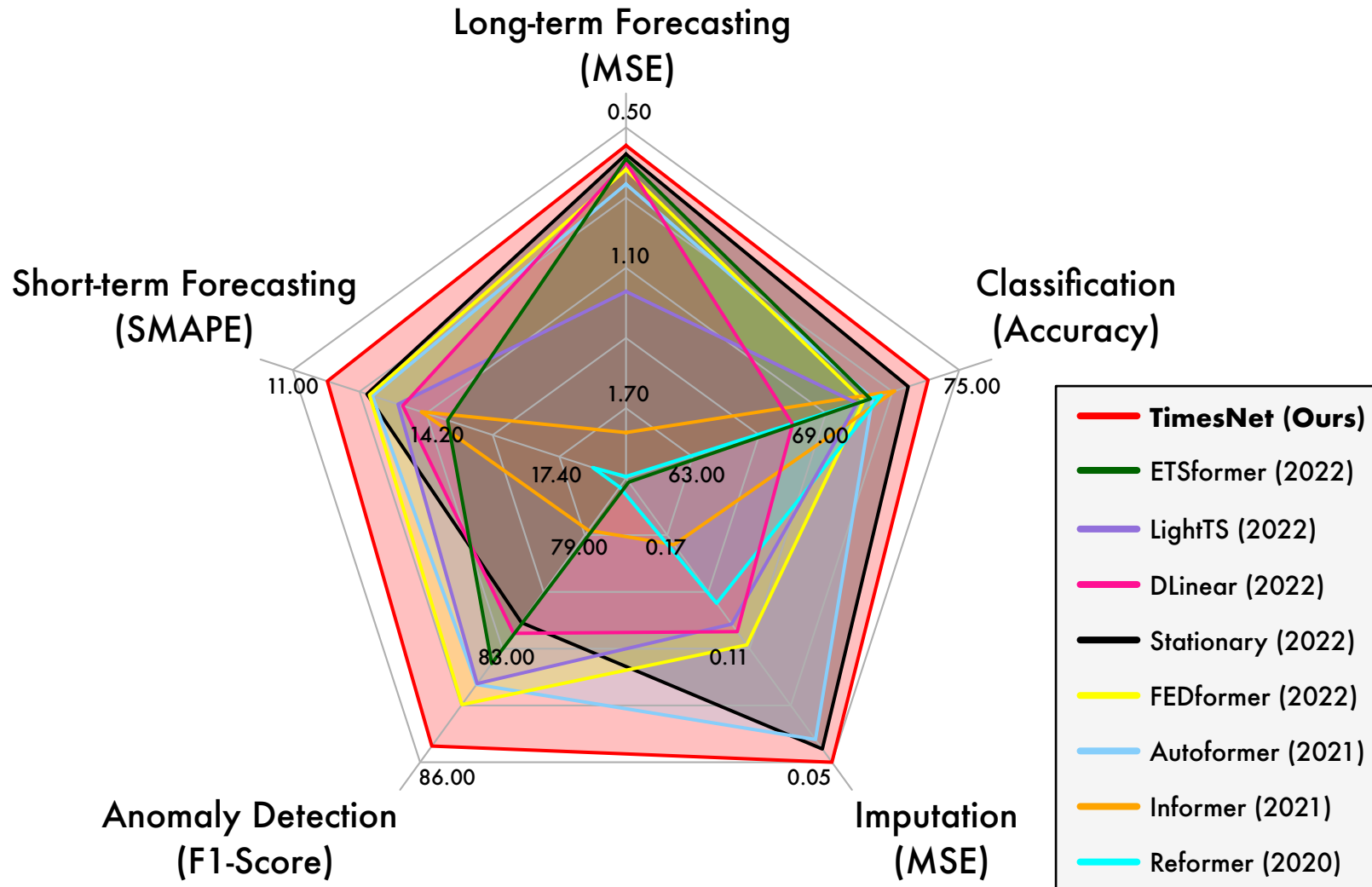
- ① 1D  $\rightarrow$  2D
- ② 2D representation learning
- ③ 2D  $\rightarrow$  1D

# Experiment: Overall

Tasks	Benchmarks
Forecasting	<b>Long-term:</b> ETT (4 subsets), Electricity, Traffic, Weather, Exchange, ILI
	<b>Short-term:</b> 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

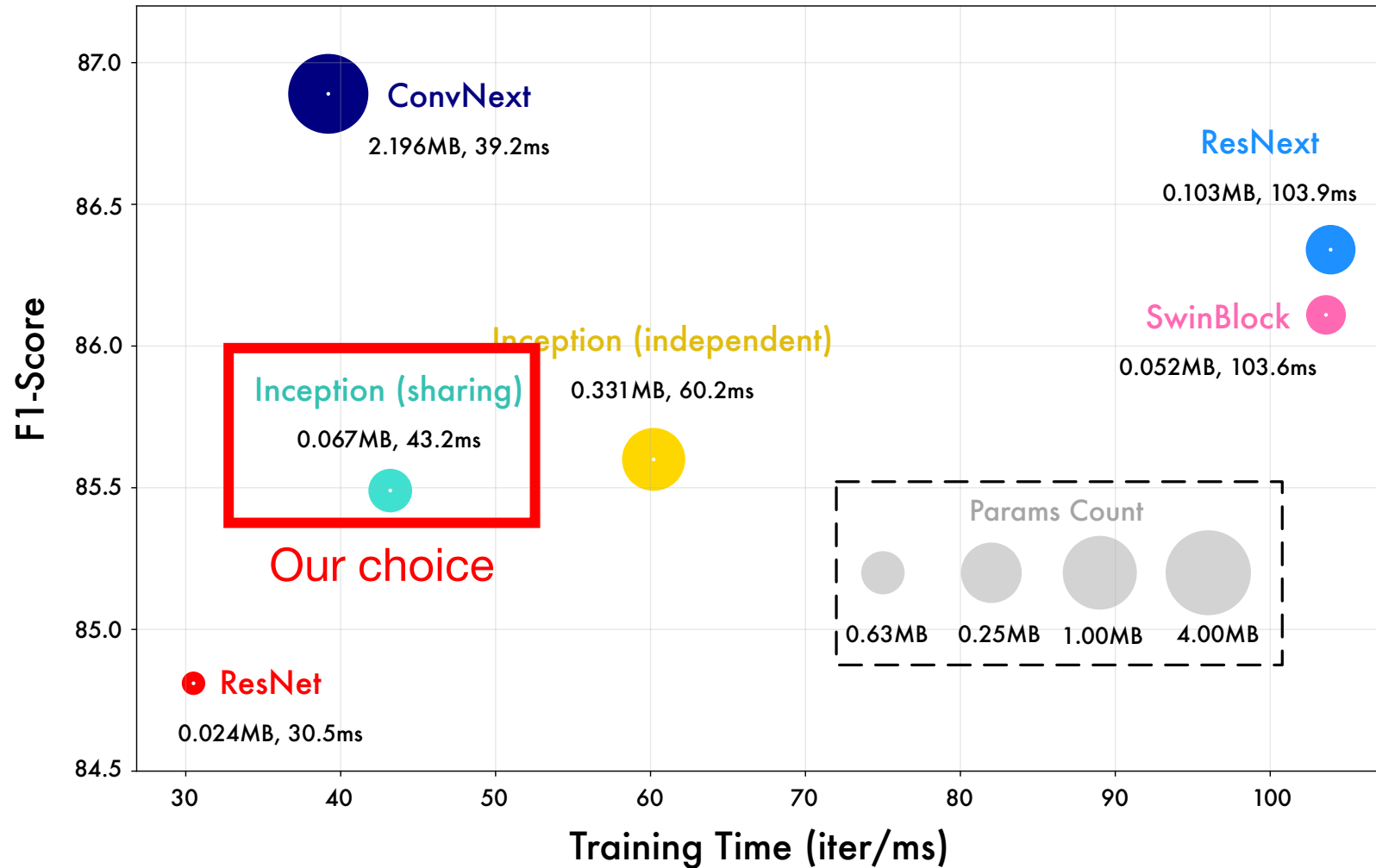
# Experiment: Overall



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 514 forks, 3k stars, and 41 watchers. The main branch is 'main'. The repository contains a list of files and folders, including 'data\_provider', 'exp', 'layers', 'models', 'pic', 'scripts', 'tutorial', 'utils', '.gitignore', 'LICENSE', 'README.md', 'requirements.txt', and 'run.py'. The 'tutorial' folder is highlighted. The repository description is 'A Library for Advanced Deep Time Series Models.' and it is licensed under MIT. The repository has 22 contributors and 8 additional contributors. The languages section shows a bar chart with Python, C++, and other languages.

thuml / Time-Series-Library

Code Issues 13 Pull requests 3 Discussions Actions Projects Wiki Security 1 Insights Settings

Time-Series-Library Public

Edit Pins Watch 41 Fork 514 Starred 3k

main 1 Branch 0 Tags

Go to file Add file Code

wuhaixu2016 Merge pull request #295 from Mico3o/main 160e497 · 3 days ago 160 Commits

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

README MIT license

## Time Series Library (TSlib)

TSlib is an open-source library for deep learning researchers, especially for deep time series analysis.

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

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<sup>✉</sup>

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

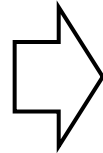
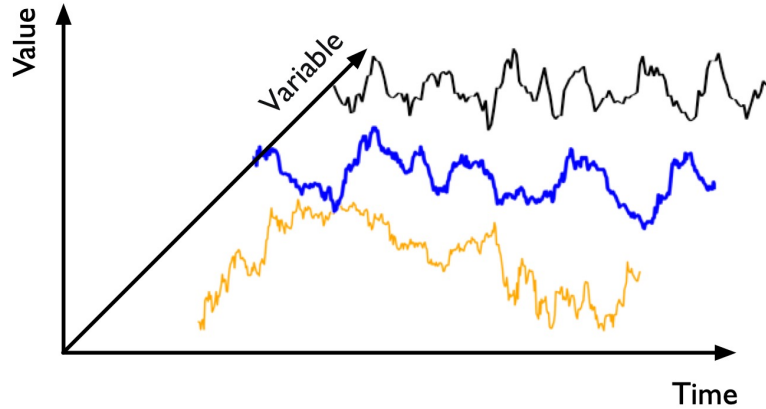


Jianmin Wang

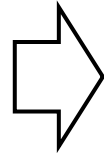


Mingsheng Long

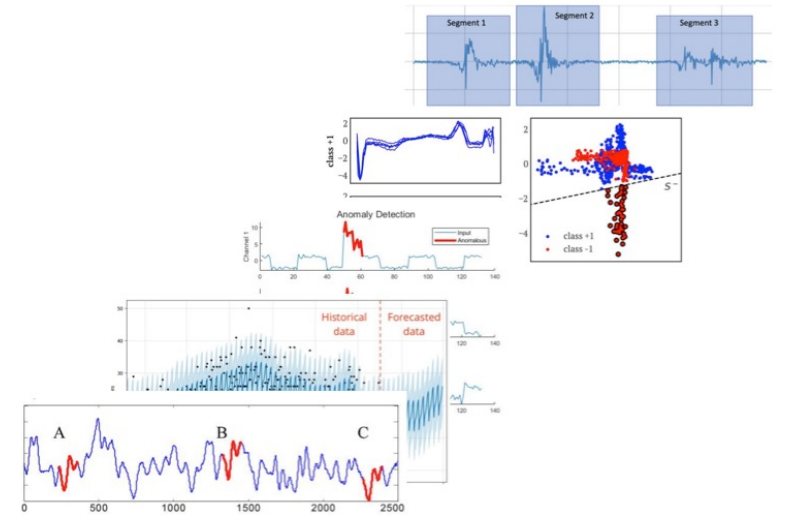
# Time Series Pre-training



Pre-training



Fine-tuning



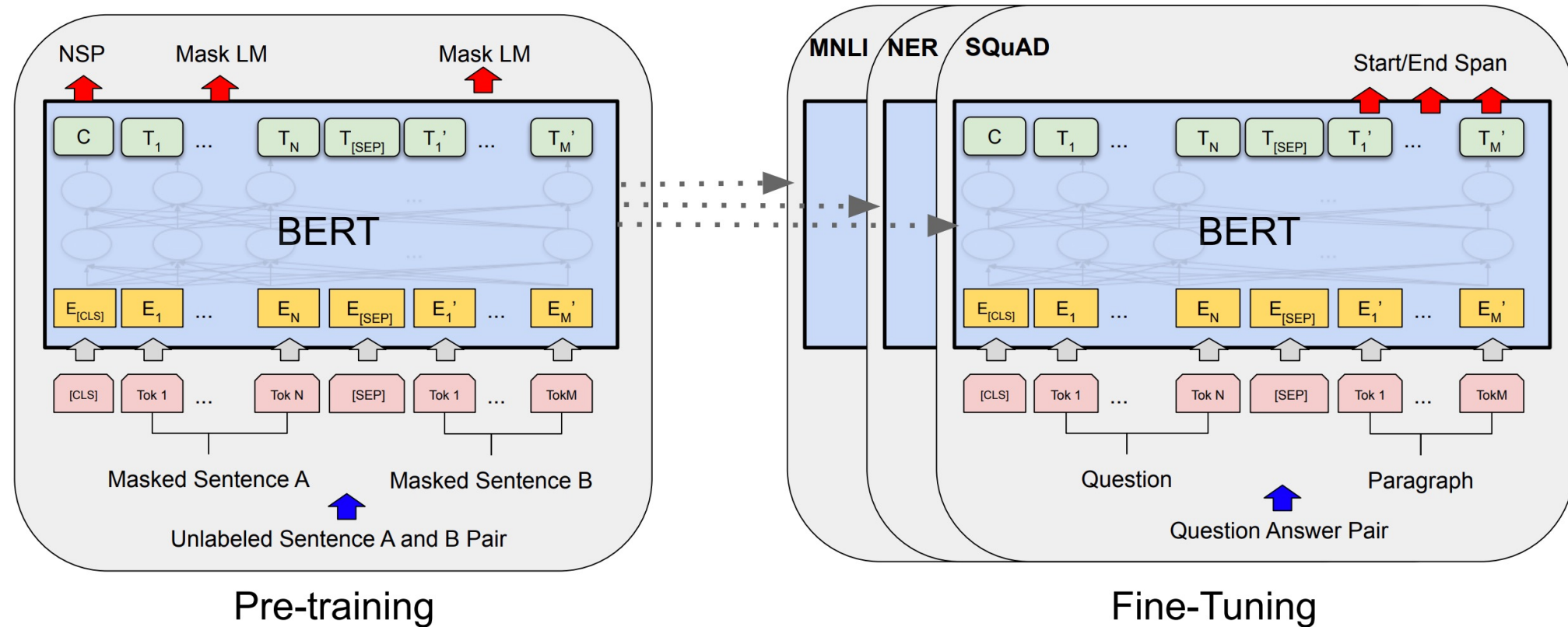
Large-scale time series data

Diversified time series analysis tasks

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

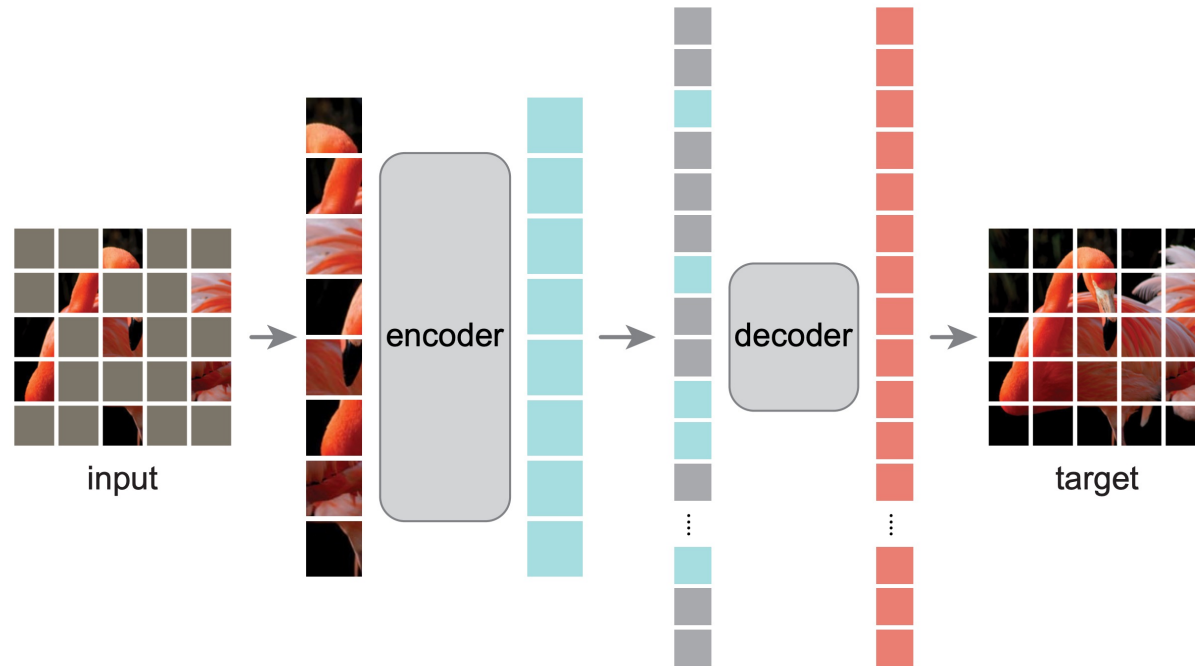
# Masked Modeling in NLP

Random mask a portion of words.

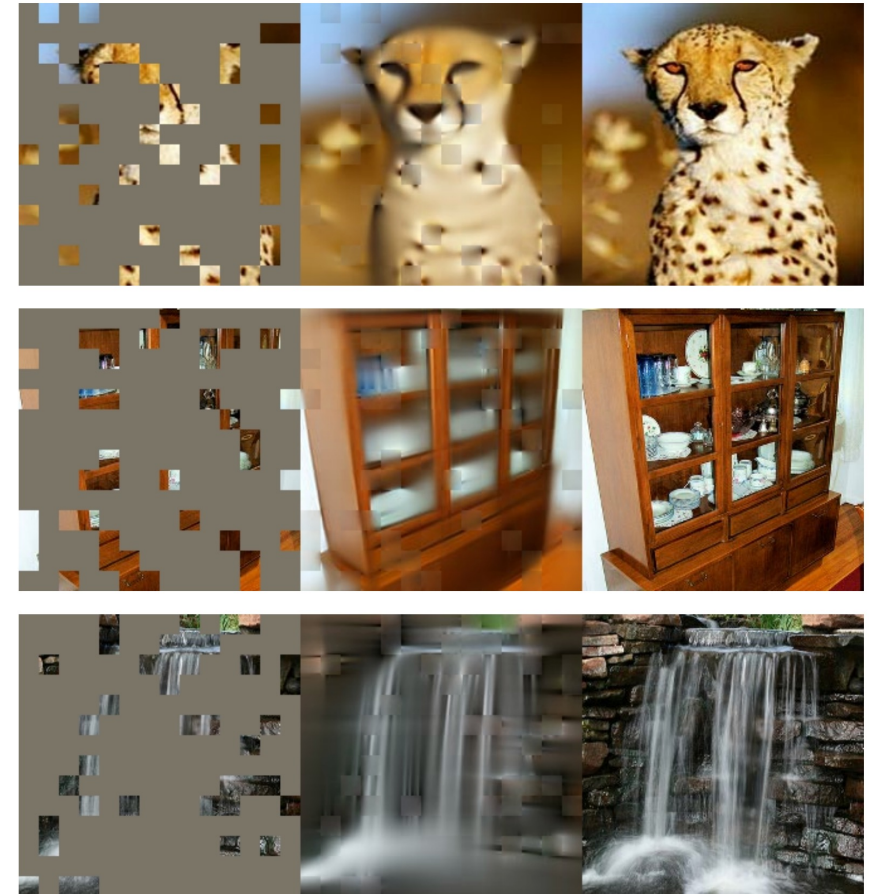




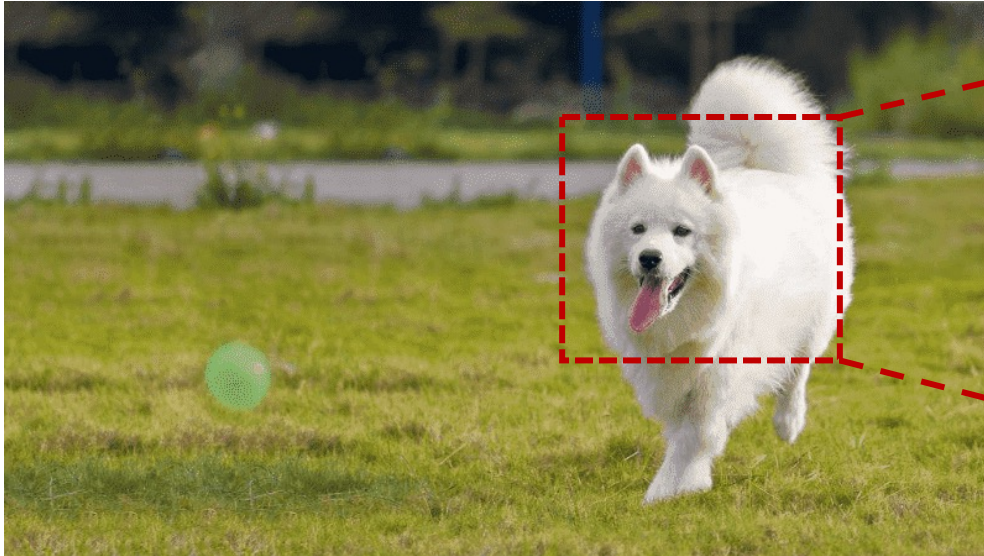
# Masked Modeling in CV



Random mask a portion of patches.

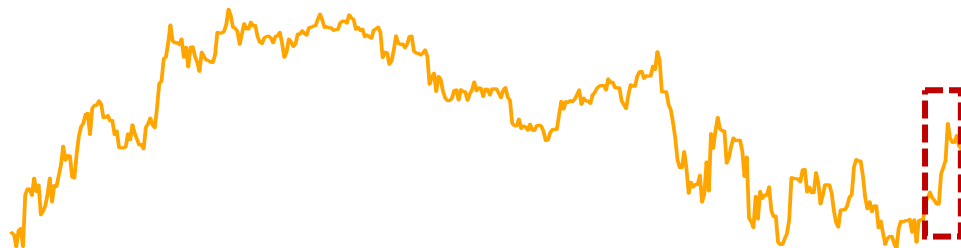


# Differences among Image, Language, Time Series



TimesNet is for time series analysis.

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

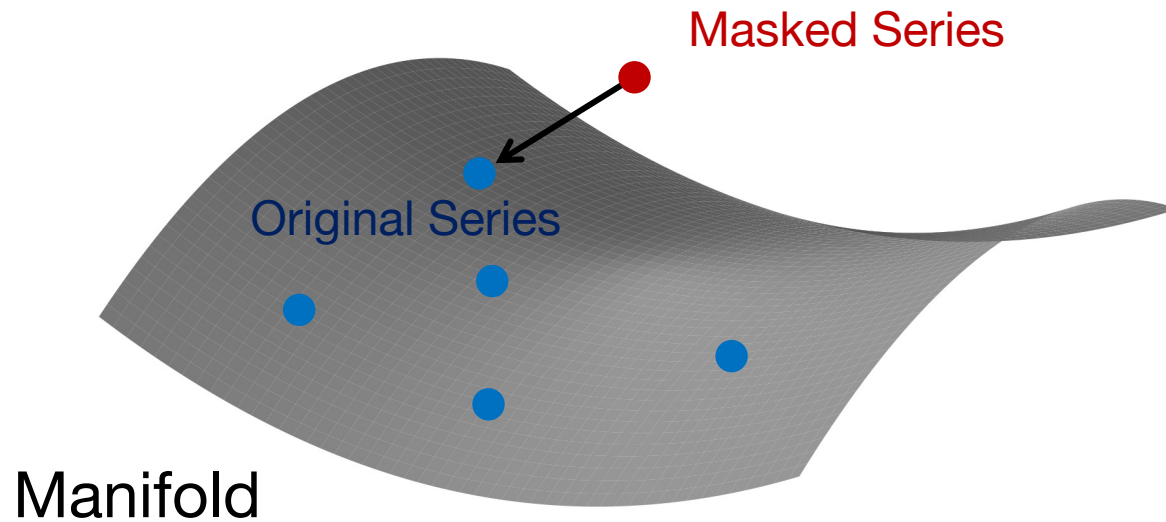


**Each time point only saves some scalars.**

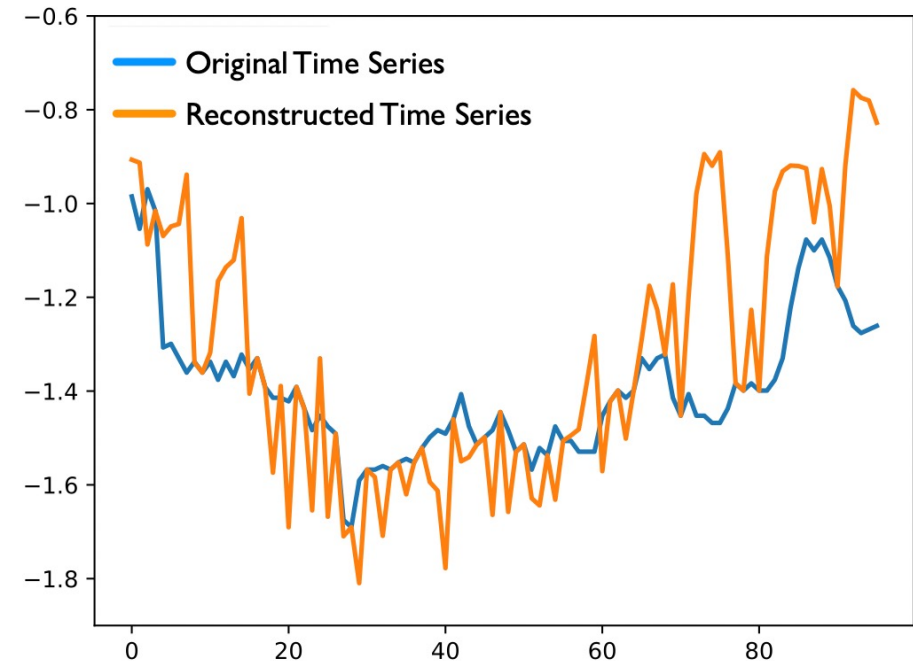


WIKIPEDIA  
The Free Encyclopedia

# Canonical Masked Modeling in Time Series



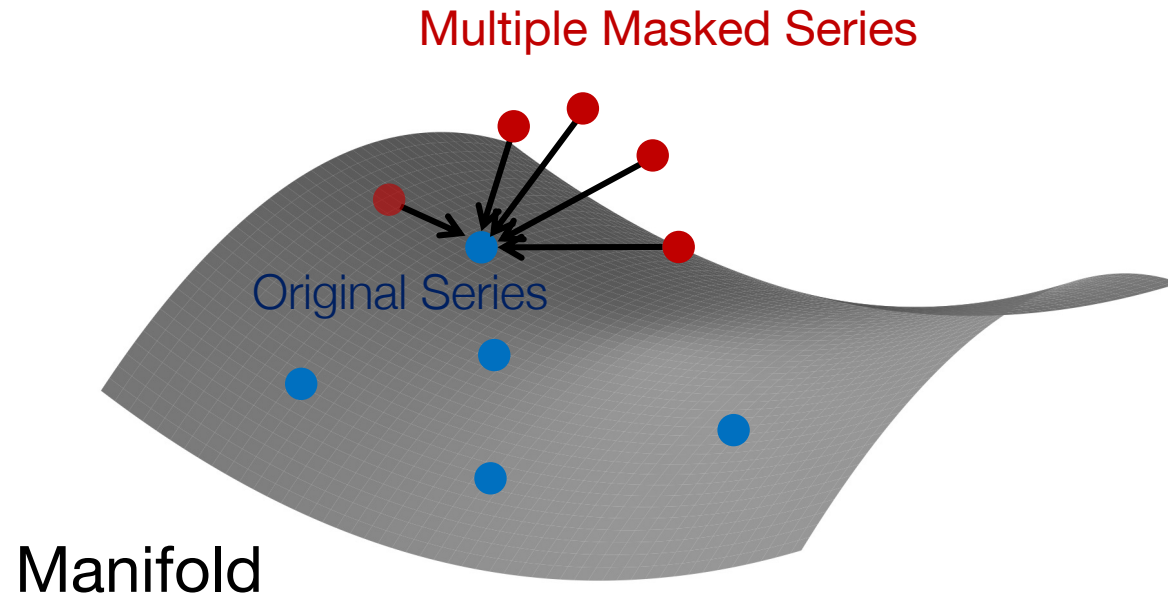
## Hard to Reconstruct



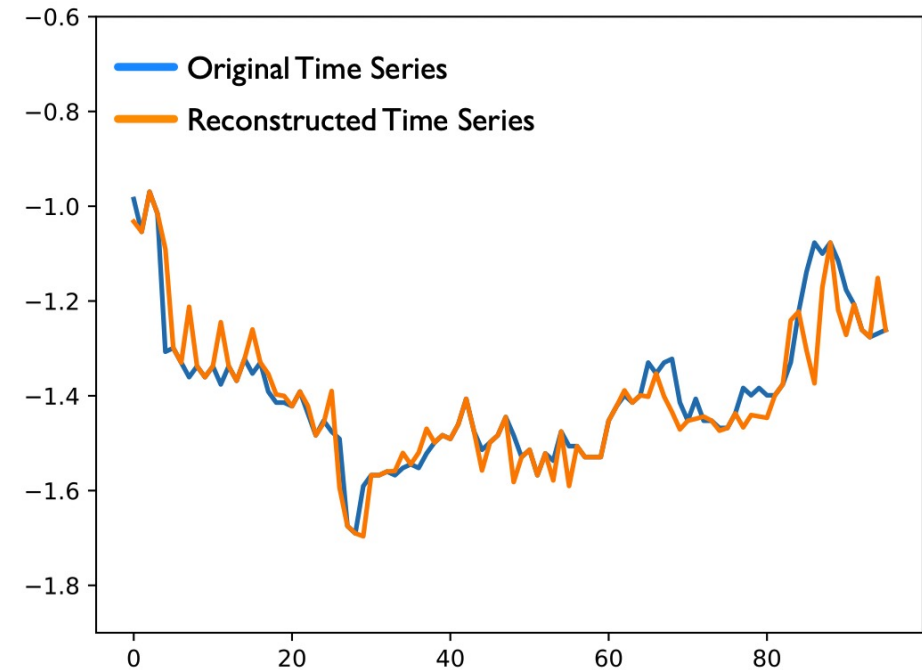
## ✓ Direct Reconstruction

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

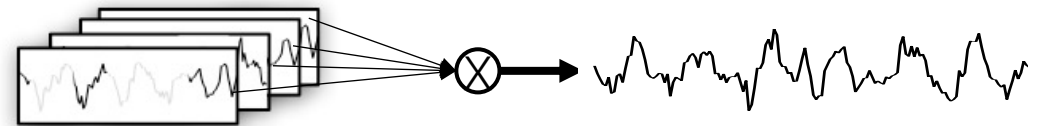
## Canonical



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

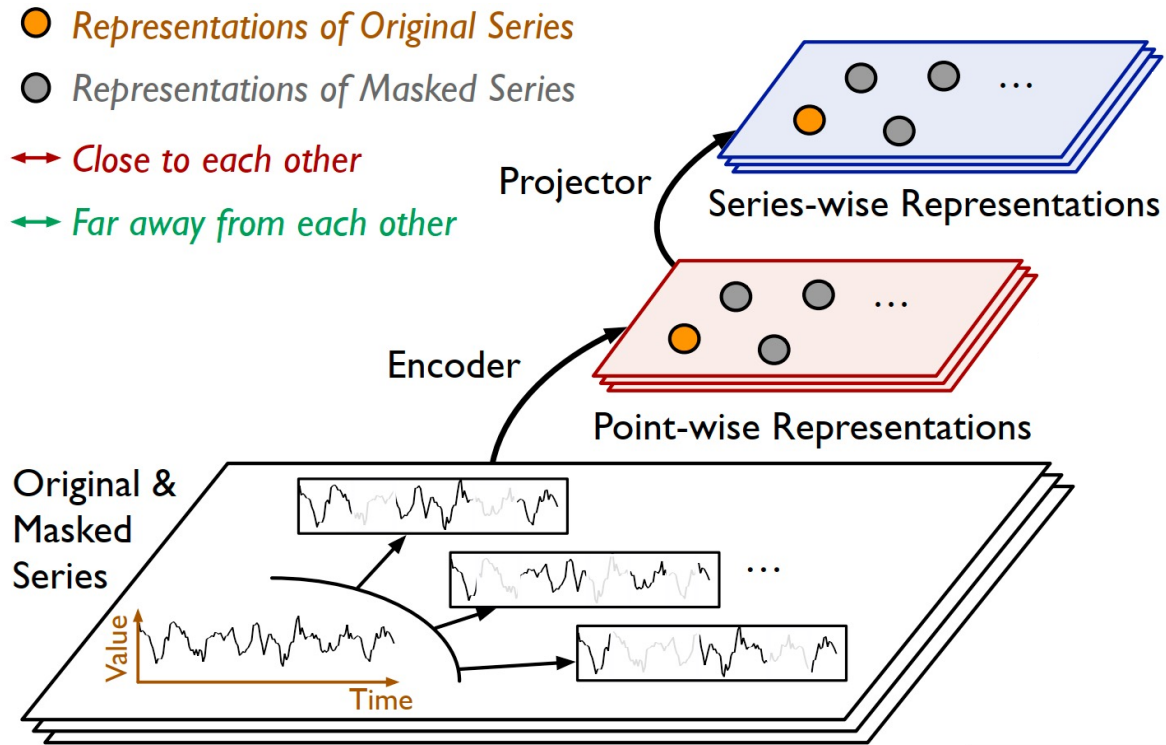
**VS**

## Neighborhood Aggregation



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

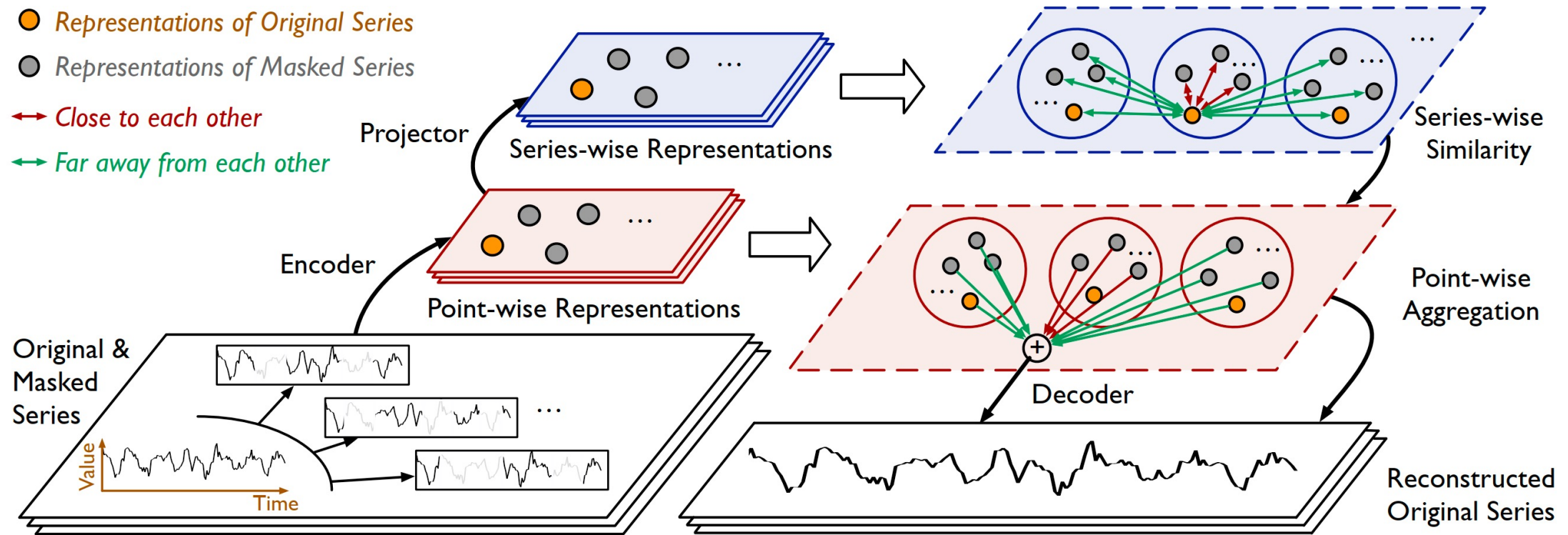
# Overall design of SimMTM



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.

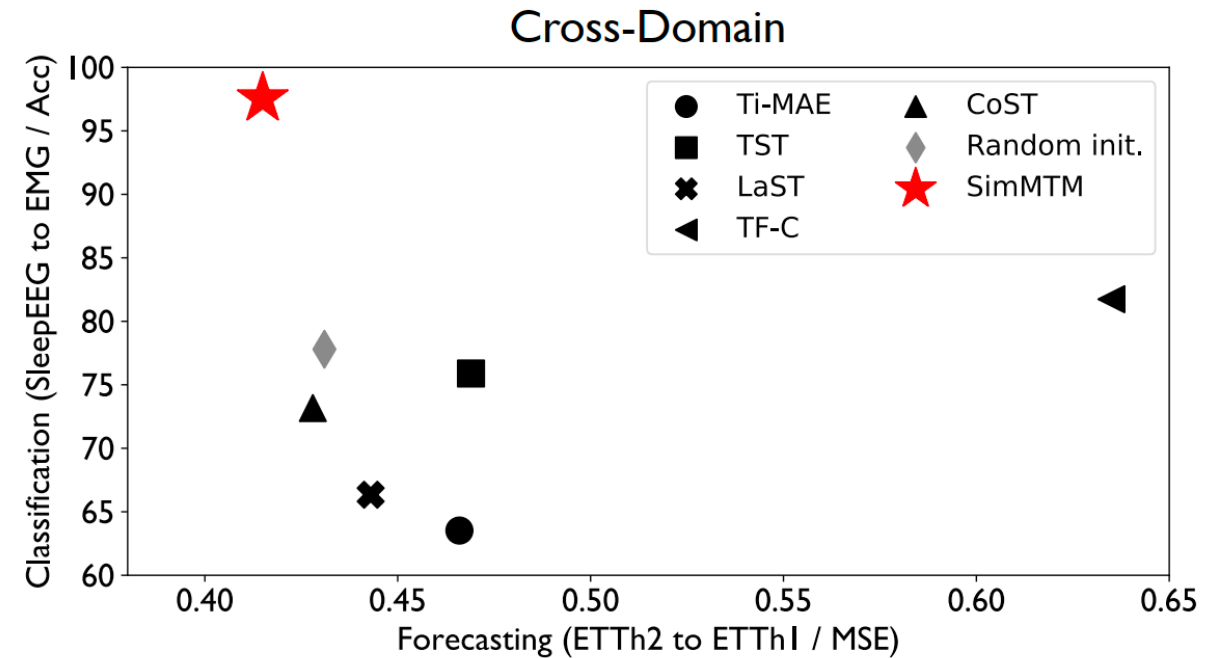
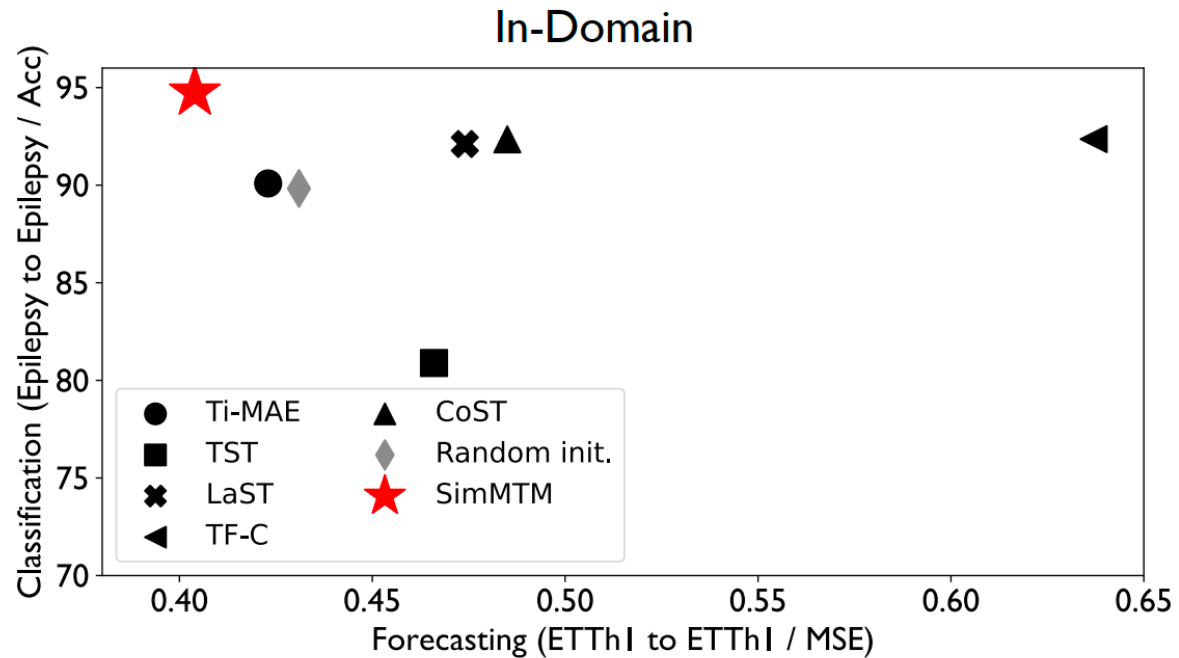
# Experiment: Overall

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	<b>0.927</b>	<b>0.761</b>	<b>3.498</b>	<b>1.487</b>	<b>0.809</b>	<b>0.663</b>	<b>1.322</b>	<b>0.808</b>	
Autoformer [47]	0.573	0.573	0.550	0.559	0.615	0.528	0.324	0.368	↑
+ SimMTM	<b>0.561</b>	<b>0.568</b>	<b>0.543</b>	<b>0.555</b>	<b>0.553</b>	<b>0.505</b>	<b>0.315</b>	<b>0.360</b>	
NS Transformer [24]	0.570	0.537	0.526	0.516	0.481	0.456	0.306	0.347	↑
+ SimMTM	<b>0.543</b>	<b>0.527</b>	<b>0.493</b>	<b>0.514</b>	<b>0.431</b>	<b>0.455</b>	<b>0.301</b>	<b>0.345</b>	
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↓	<b>0.341</b>	0.379	0.258	0.318↓	
+ SimMTM	<b>0.409</b>	<b>0.428</b>	<b>0.329</b>	0.379	0.348	<b>0.378</b>	<b>0.254</b>	<b>0.313</b>	

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

# Open Source

The screenshot displays the GitHub repository page for 'SimMTM' by user 'thuml'. The repository is public and has 29 stars, 5 watchers, and 2 forks. The commit history shows a recent commit by 'dongjiaxiang' to 'adjust readme'. The README section features the title 'SimMTM (NeurIPS 2023)' and a link to the paper: 'SimMTM: A Simple Pre-Training Framework for Masked Time-Series Modeling'. Below the title is an 'Architecture' diagram (Figure 1) illustrating the process from 'Original & Masked Series' through an 'Encoder' to 'Point-wise Representations', then through a 'Projector' to 'Series-wise Representations', and finally through a 'Decoder' to 'Reconstructed Original Series'. The diagram also shows 'Self-supervised Pre-training' involving 'Series-wise Similarity' and 'Point-wise Aggregation'. A legend indicates that orange dots represent 'Representations of Original Series', blue dots represent 'Representations of Masked Series', red dashed lines indicate 'Close to each other', and green dashed lines indicate 'Far away from each other'. The right sidebar contains repository statistics and sections for 'About', 'Releases', 'Packages', 'Languages' (Python 96.9%, Shell 3.1%), and 'Suggested workflows'.

File	Commit	Author	Date
SimMTM_Classification	init commit		last month
SimMTM_Forecasting	init commit		last month
figs	init commit		last month
README.md	adjust readme	dongjiaxiang	last month

**SimMTM (NeurIPS 2023)**

This is the codebase for the paper: [SimMTM: A Simple Pre-Training Framework for Masked Time-Series Modeling](https://arxiv.org/abs/2302.00861)

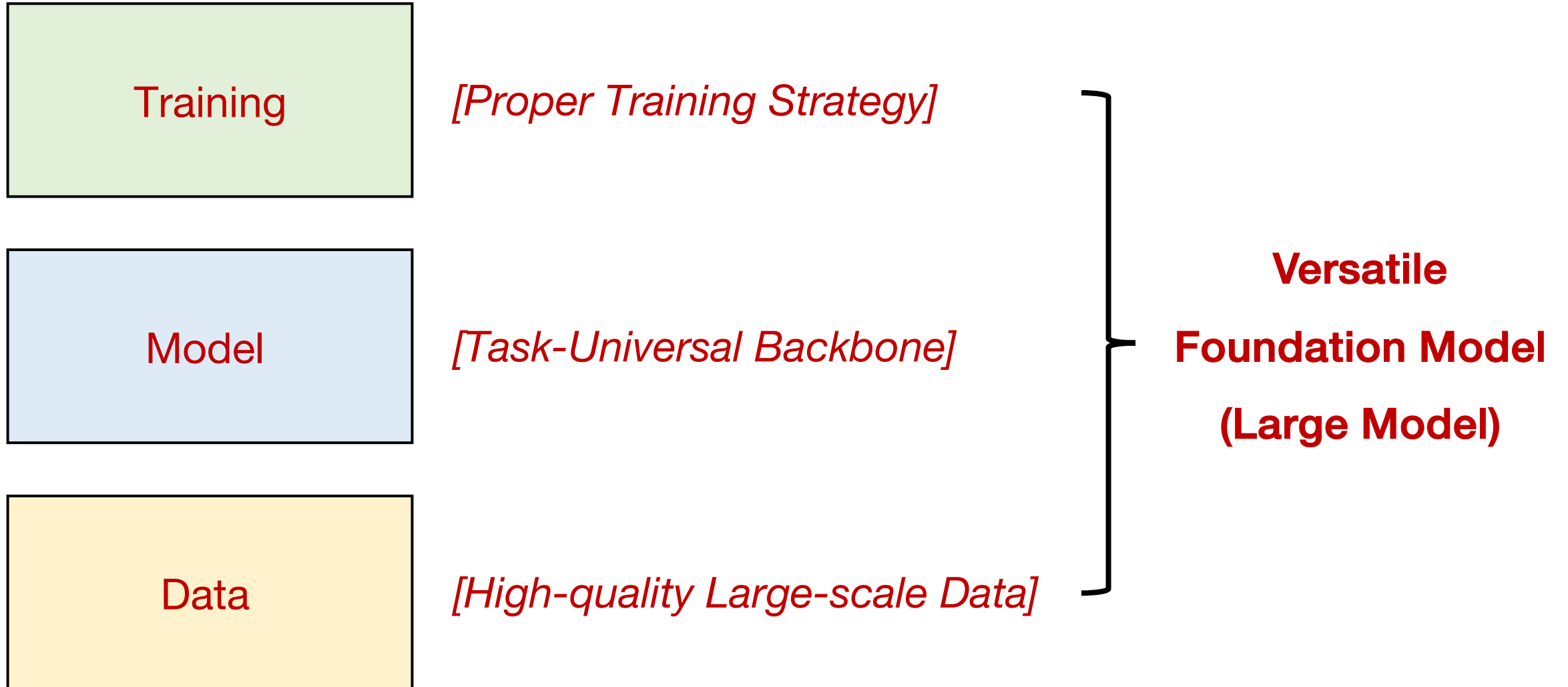
**Architecture**

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.

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

# 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