ARISE
Automated Reliable Intelligent Software Engineering

香港中文大學
CUHK

Research Track

Maat: Performance Metric Anomaly Anticipation for Cloud Services with Conditional Diffusion

Cheryl Lee[*], Tianyi Yang[*], Zhuangbin Chen[†], Yuxin Su[†], Michael R. Lyu[*]

[*]The Chinese University of Hong Kong
[†]Sun Yat-sen University

Sep, 2023

Table of Contents

01 INTRODUCTION

Motivation

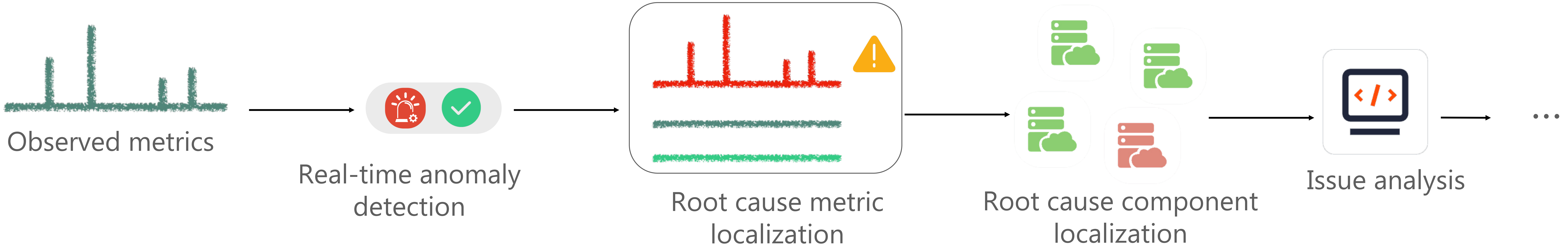
Twitter Back After Two-Hour Outage Affected Tweets

Facebook Lost About $65 Million During Hours-Long Outage

YouTube App Down on iOS Devices

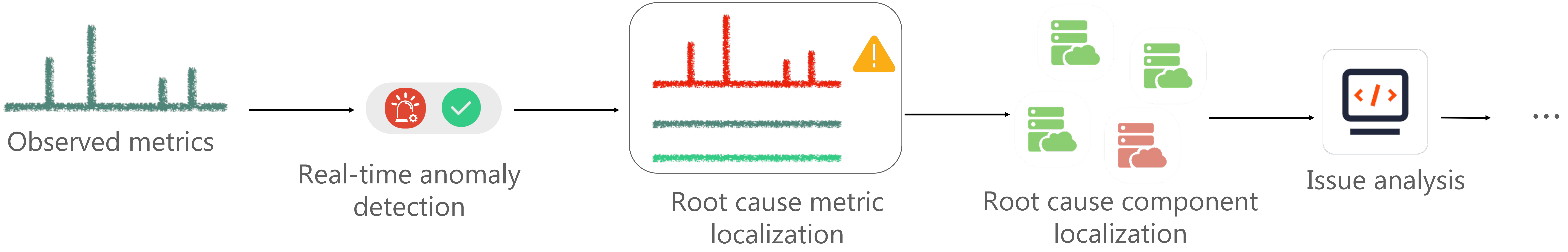
Amazon’s One Hour of Downtime on Prime Day May Have Cost It up to $100 Million in Lost Sales

Motivation

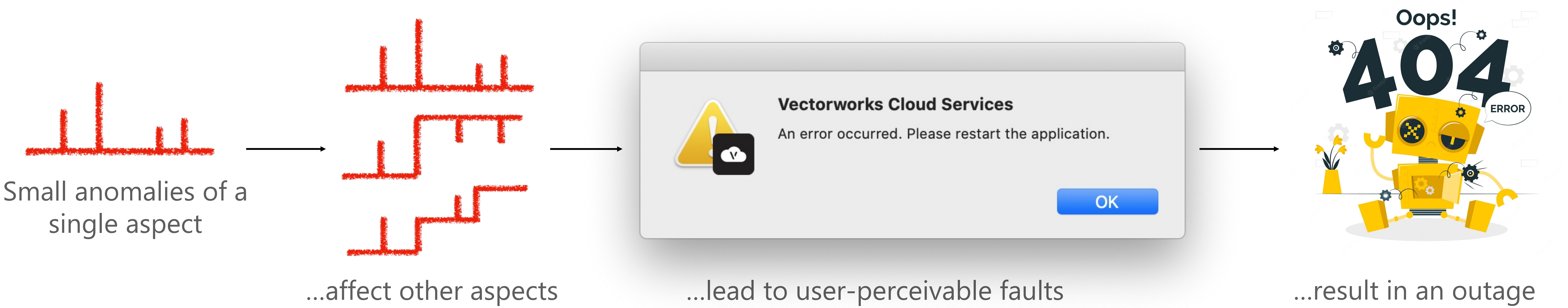


Observed metrics
Real-time anomaly detection
Root cause metric localization
Root cause component localization
Issue analysis
…

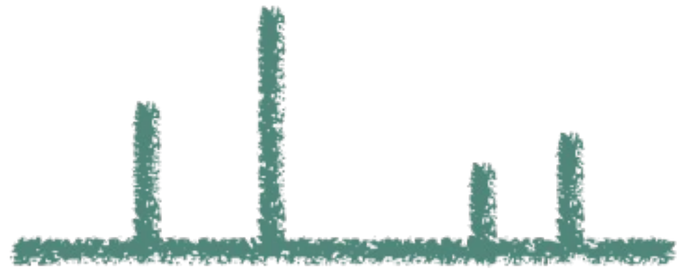
Motivation



At the same time...



Motivation



Observed metrics

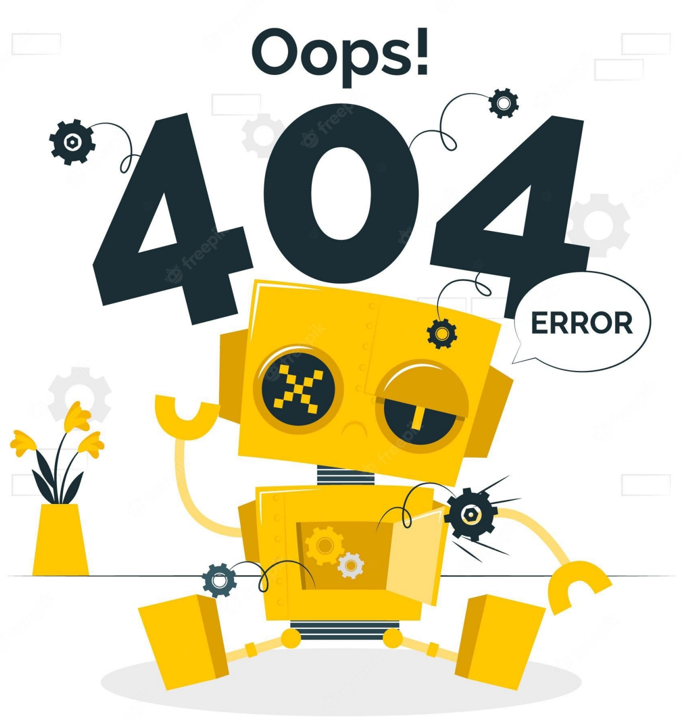
At the same time

Small anomalies of a single aspect

Too Late!

Issue analysis

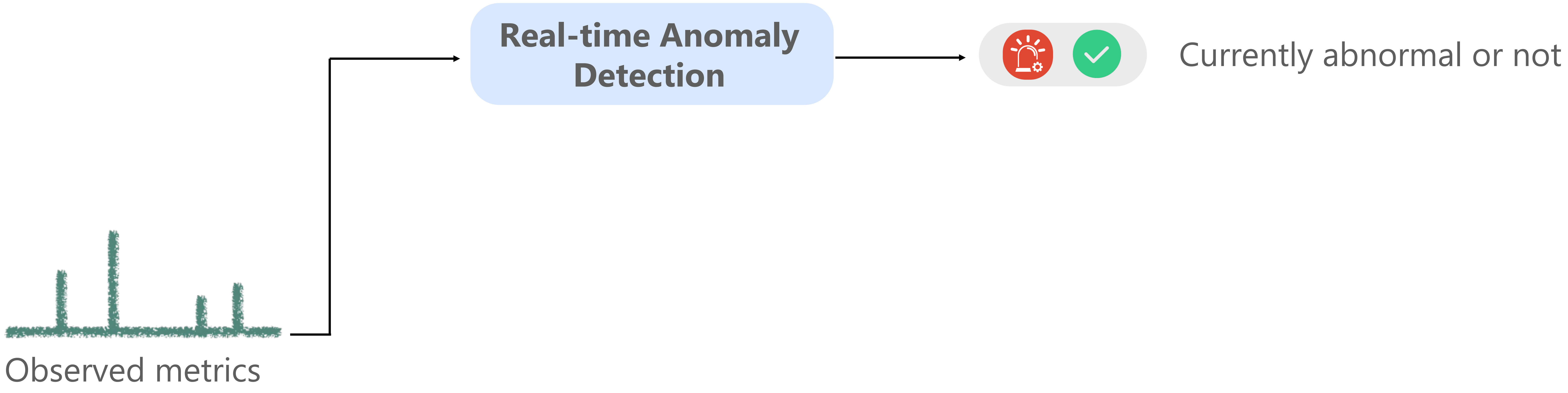
...



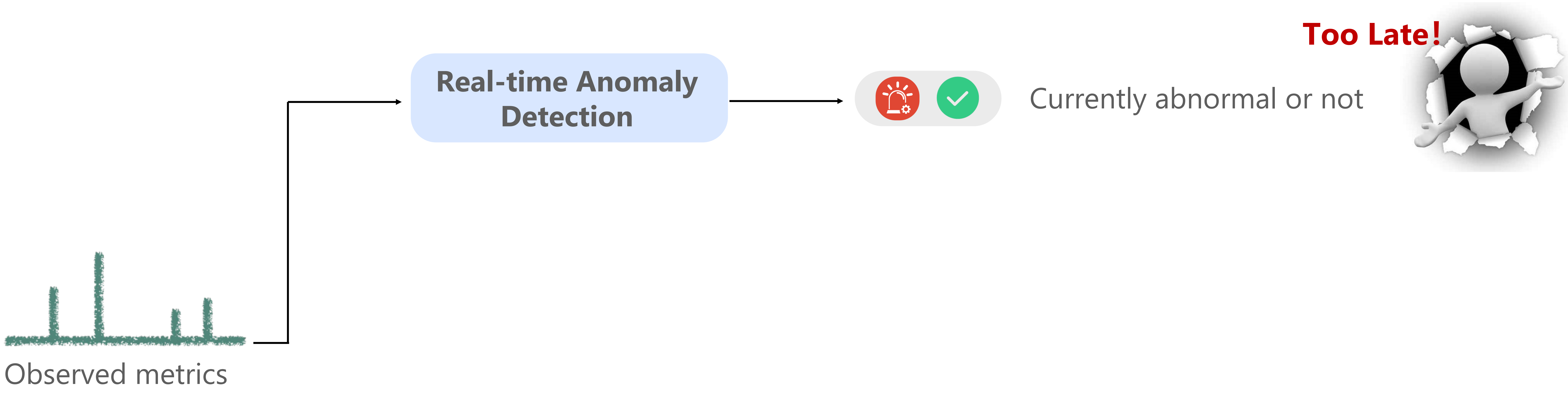
...result in an outage

02 PARADIGM

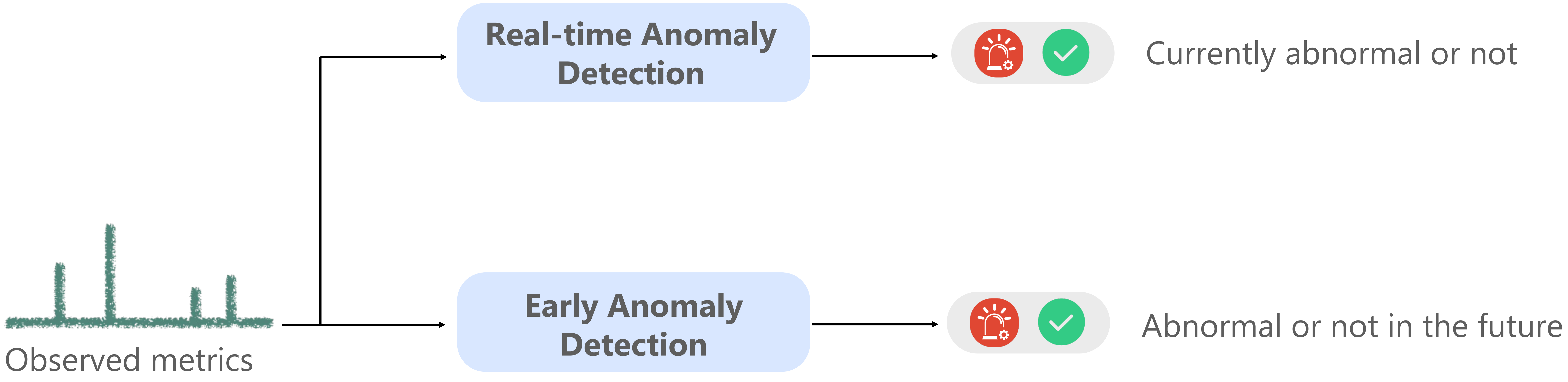
Problem Formulation: Anomaly Anticipation



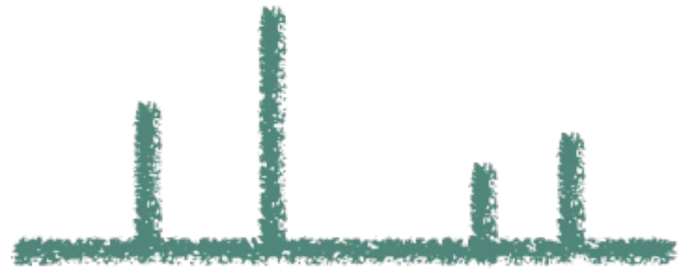
Problem Formulation: Anomaly Anticipation



Problem Formulation: Anomaly Anticipation



Problem Formulation: Anomaly Anticipation



Observed metrics

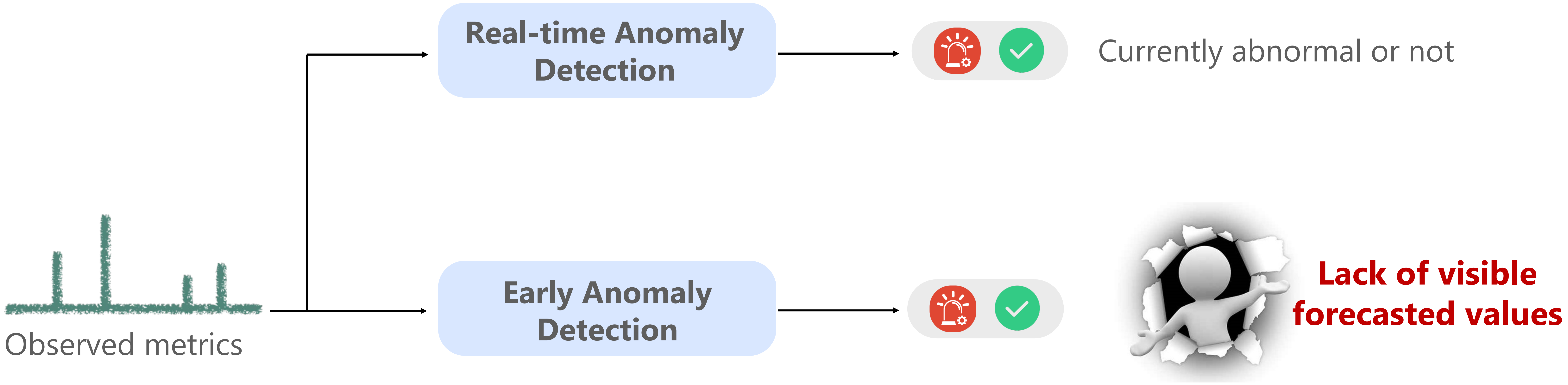
Real-time Anomaly Detection

Currently abnormal or not

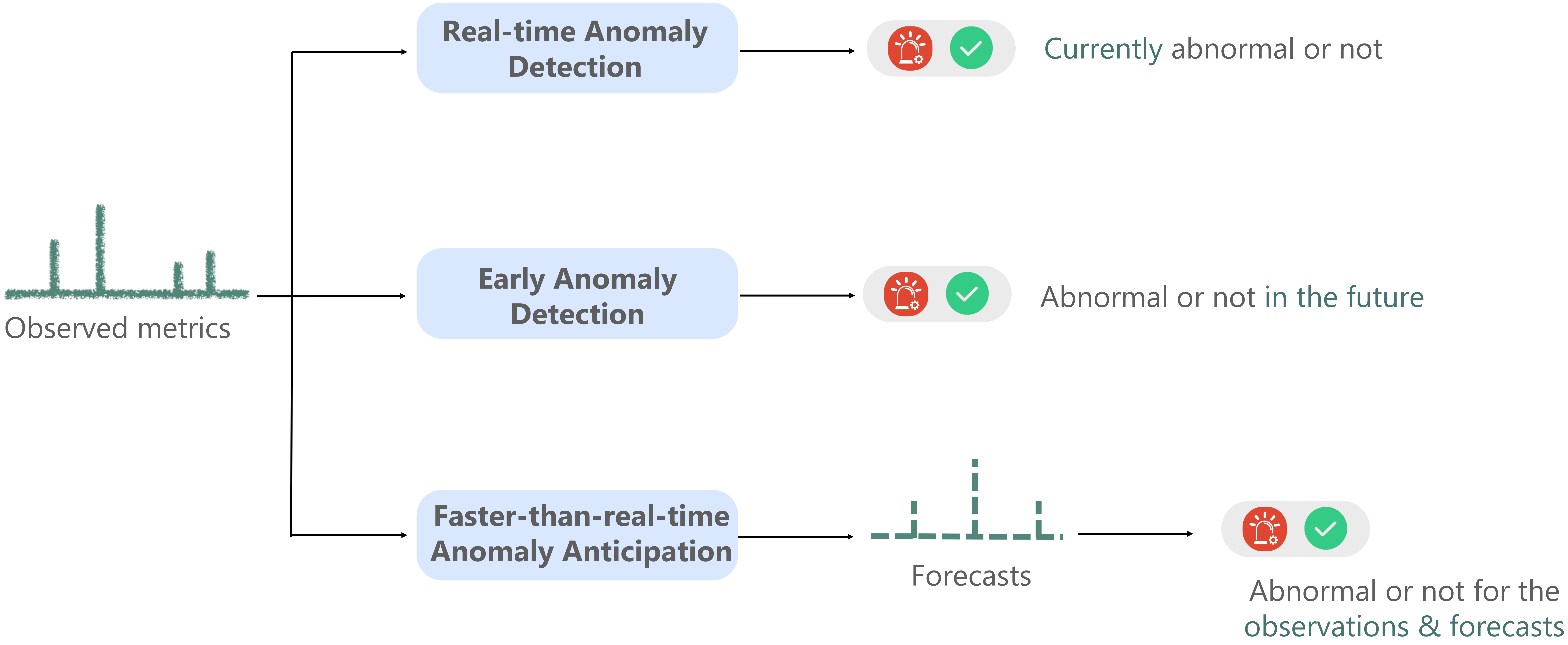
Early Anomaly Detection

Lack of high-quality labels

Problem Formulation: Anomaly Anticipation

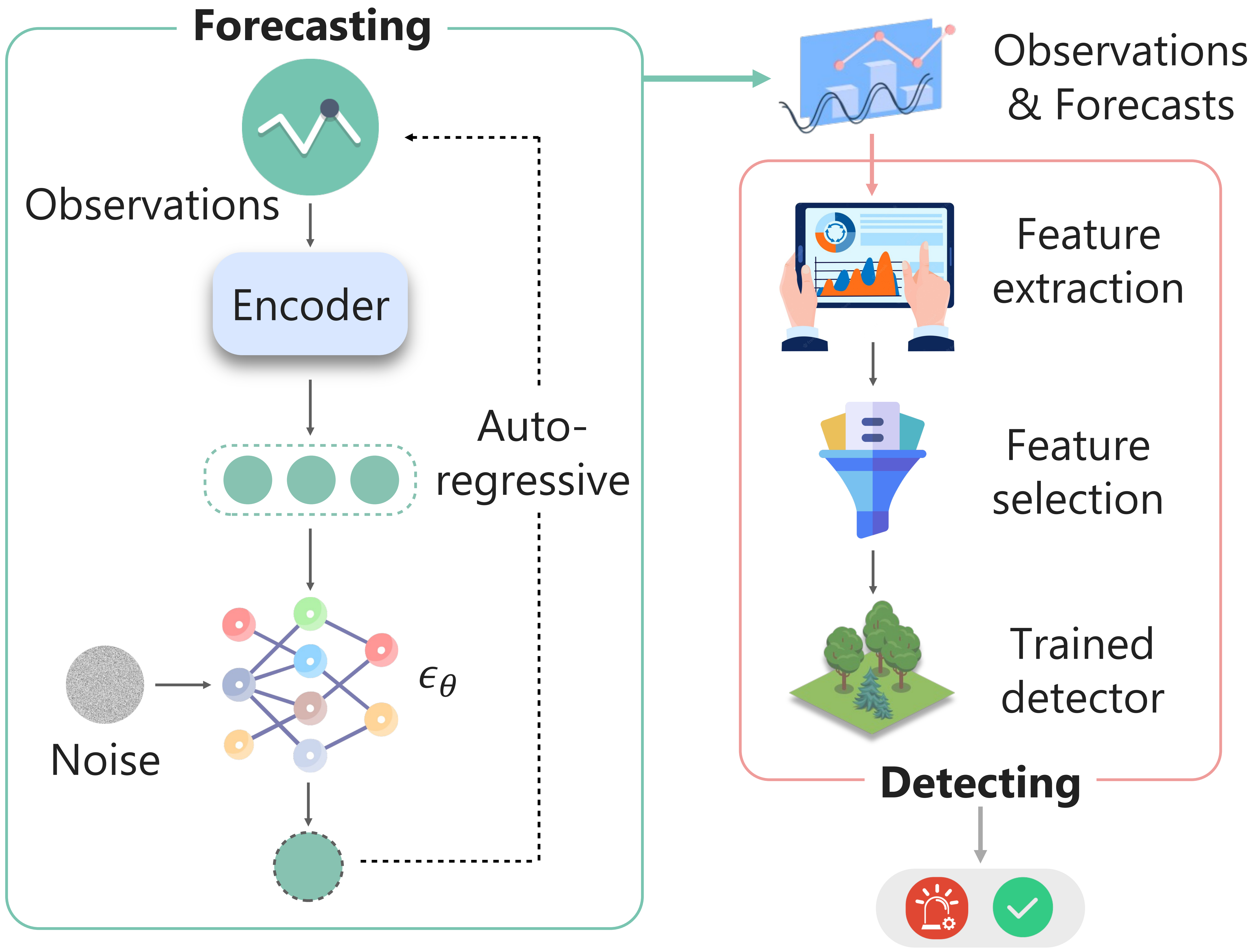


Problem Formulation: Anomaly Anticipation



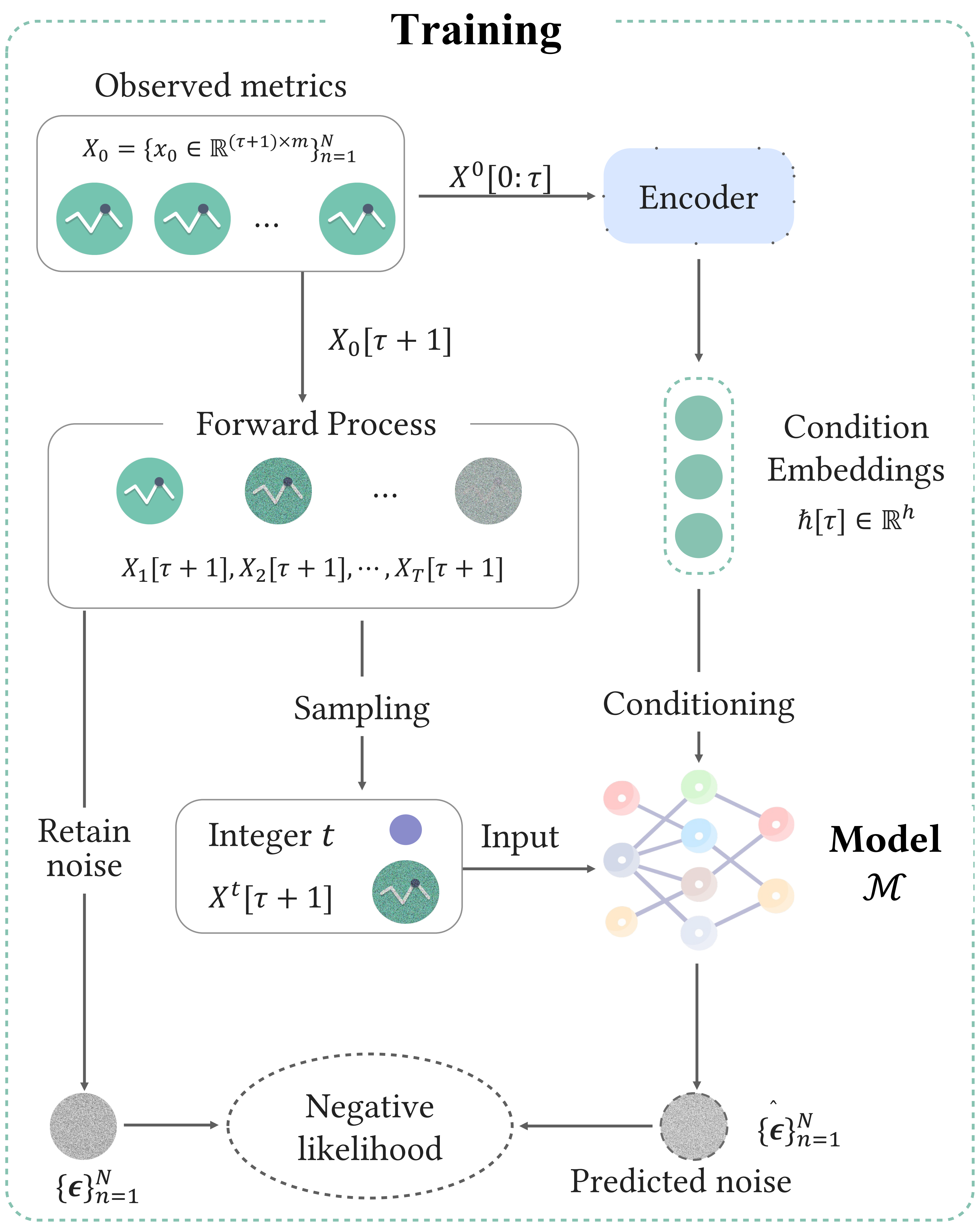
03 MOTHODOLOGY

Overview

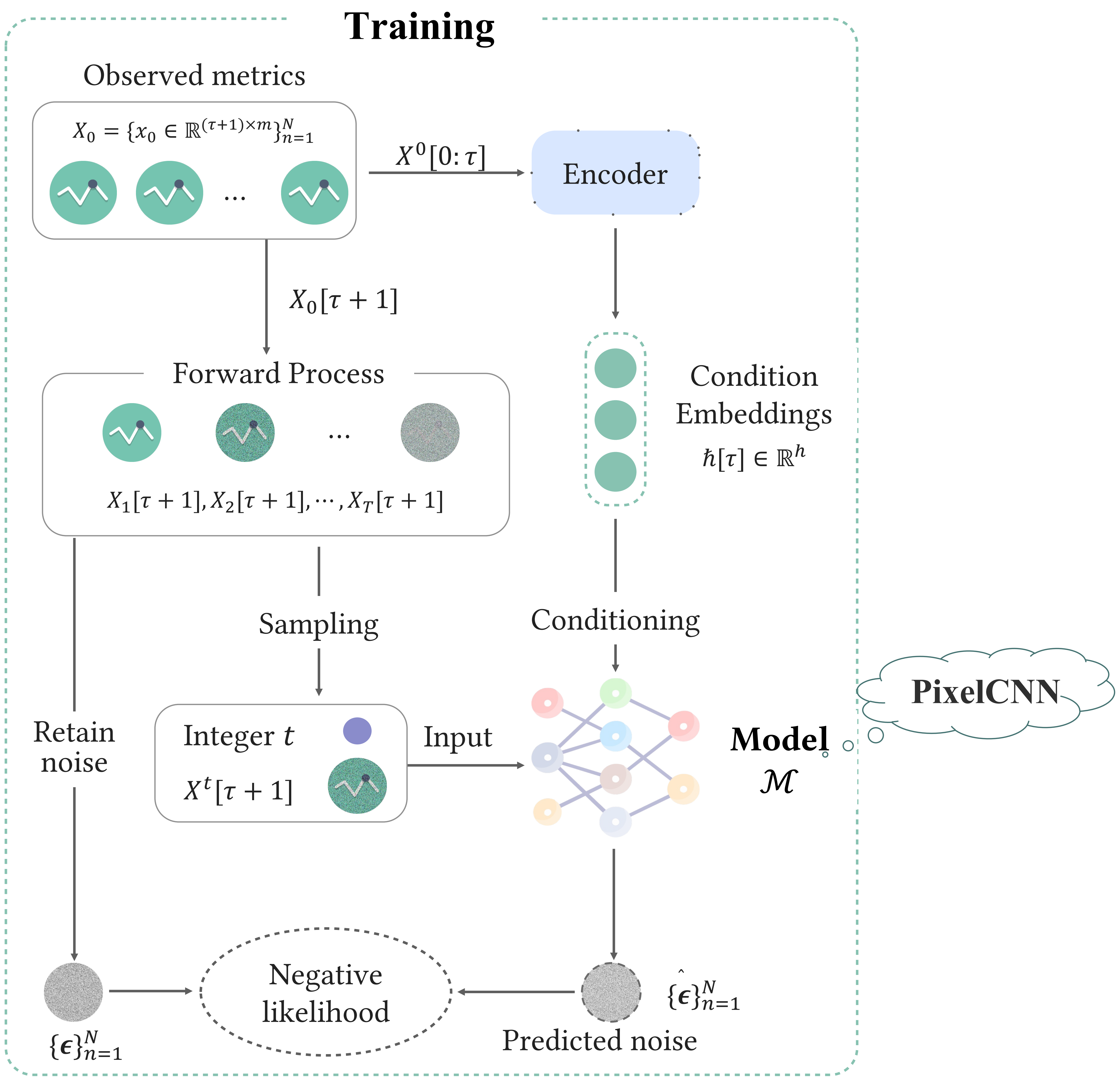


Forecasting
Observations
Encoder
Auto-
regressive
Noise
ϵ_θ
Observations
& Forecasts
Feature
extraction
Feature
selection
Trained
detector
Detecting

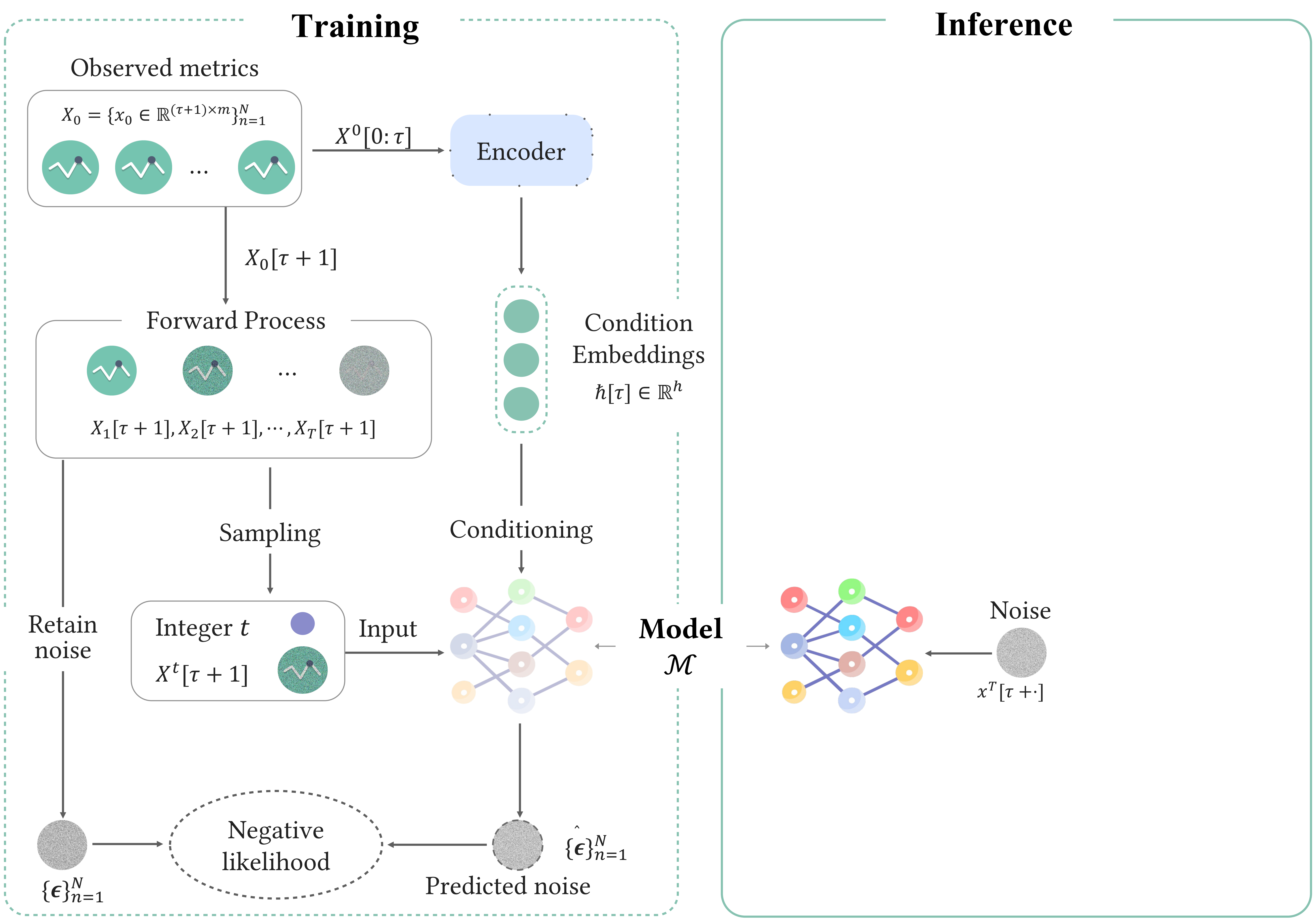
1 Forecasting



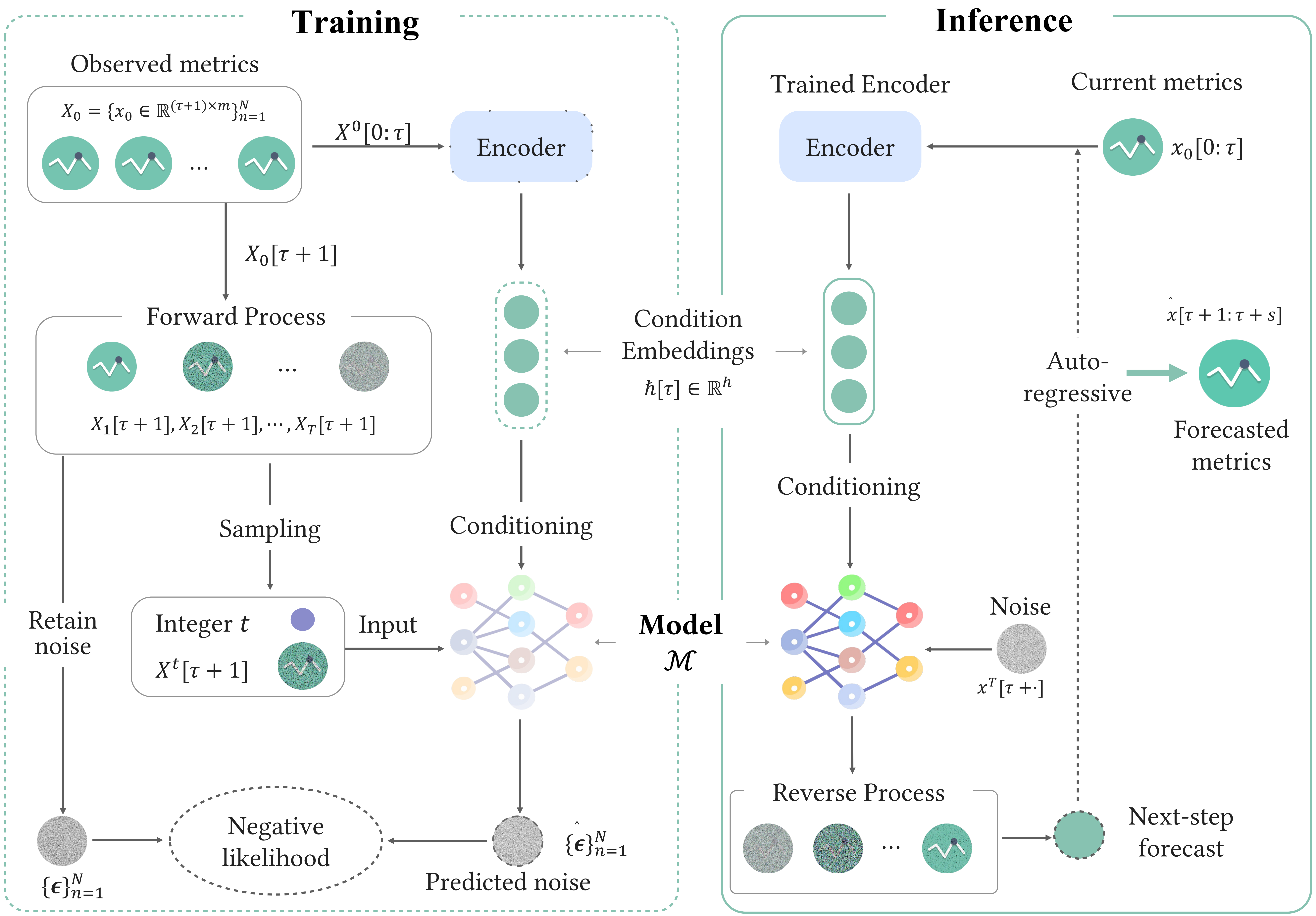
1 Forecasting



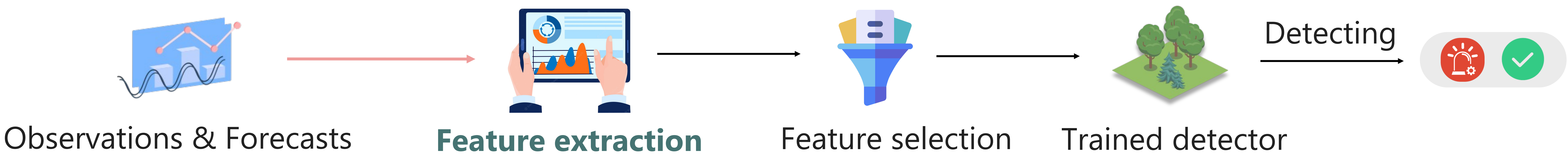
1 Forecasting



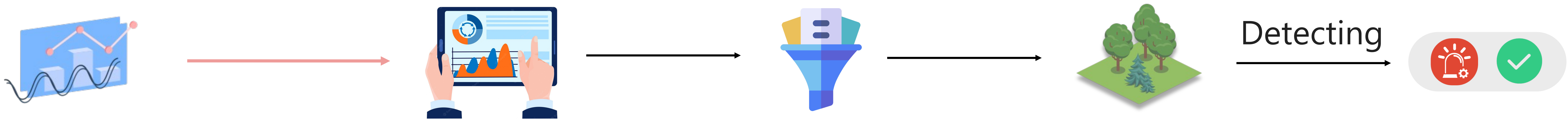
1 Forecasting



2 Detecting

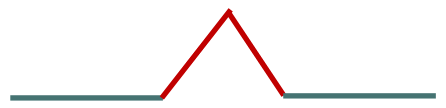

Observations & Forecasts
Feature extraction
Feature selection
Trained detector
Detecting

2 Detecting

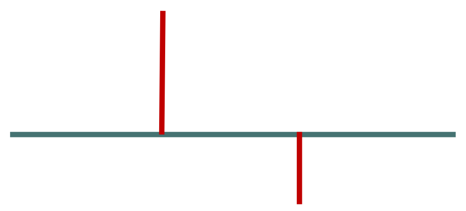


Observations & Forecasts | **Feature extraction** | Feature selection | Trained detector

Point-level

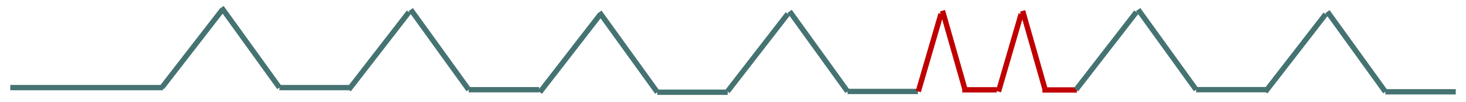

spike


jitters


peak

Frequency

Trend


period anomaly

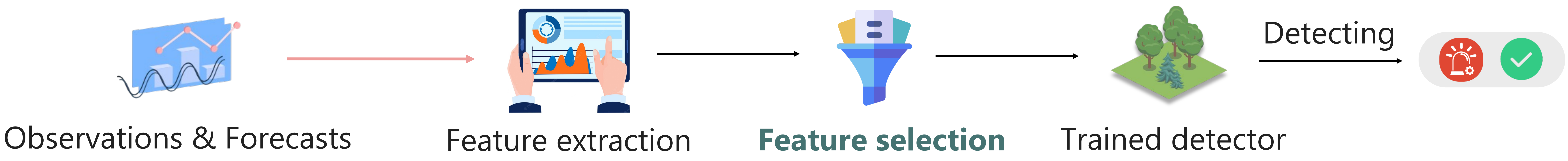
Temporal dependencies


distribution anomaly

Distribution

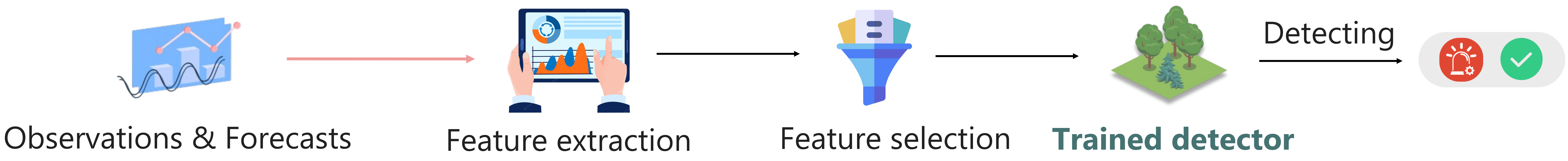
…

2 Detecting



Use Xgboost to calculate the importance score of each feature on an annotated validation set.

2 Detecting

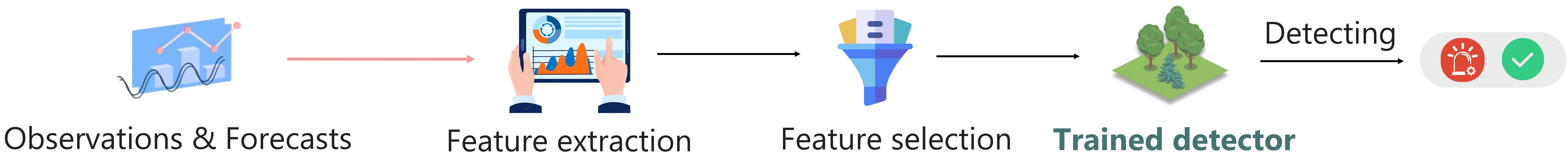


Algorithm 1: Incrementally training isolation forest.

Input: $X^{\text{cat}}_{[1:N]}$, γ, ψ, F_{pre} - previously trained forest
Output: A new forest F consisting of γ trees and F_{pre}

Initialize F
$i \leftarrow 1$ **while** $i \leq \gamma$ **do**
    $X' \leftarrow sample(X^{\text{cat}}_{[1:N]}, \psi)$
    $X'_{iso} \leftarrow F_{pre}(X')$ // Keep the samples "isolated" by F_{pre}
    $F \leftarrow F \cup iTree(X_{iso})$
end
return F

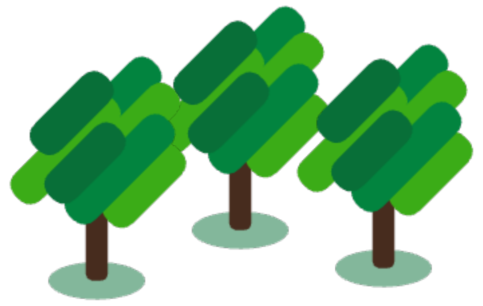
2 Detecting



Algorithm 1: Incrementally training isolation forest.

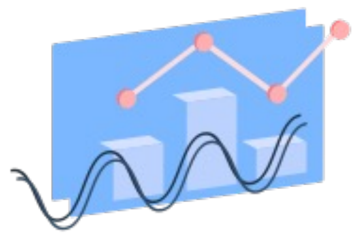
Input: $X^{\text{cat}}_{[1:N]}$, γ, ψ, F_{pre} - previously trained forest
Output: A new forest F consisting of γ trees and F_{pre}

Initialize F
$i \leftarrow 1$ **while** $i \leq \gamma$ **do**
    $X' \leftarrow sample(X^{\text{cat}}_{[1:N]}, \psi)$
    $X'_{iso} \leftarrow F_{pre}(X')$ // Keep the samples "isolated" by F_{pre}
    $F \leftarrow F \cup iTree(X_{iso})$
end
return F



Isolation trees on observations

2 Detecting



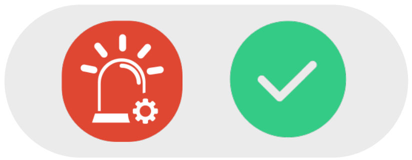
Observations & Forecasts

Feature extraction

Feature selection

Trained detector

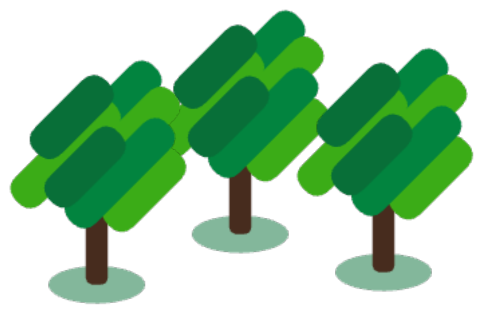
Detecting



Algorithm 1: Incrementally training isolation forest.

Input: $X^{\text{cat}}_{[1:N]}$, γ, ψ, F_{pre} - previously trained forest
Output: A new forest F consisting of γ trees and F_{pre}

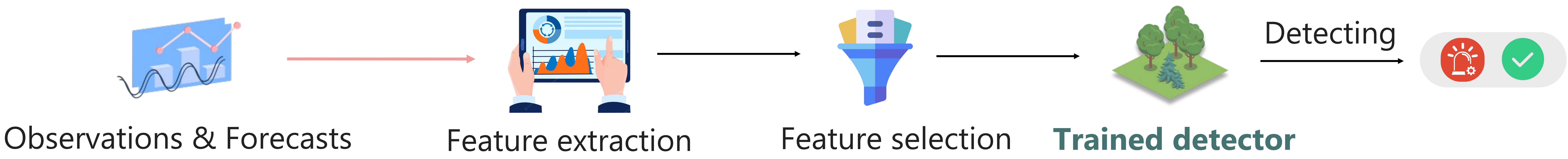
```
Initialize F
i ← 1 while i ≤ γ do
    X′ ← sample(X^cat_[1:N], ψ)
    X′_iso ← F_pre(X′) // Keep the samples "isolated" by F_pre
    F ← F ∪ iTree(X_iso)
end
return F
```



Isolation trees on observations

Incrementally learning

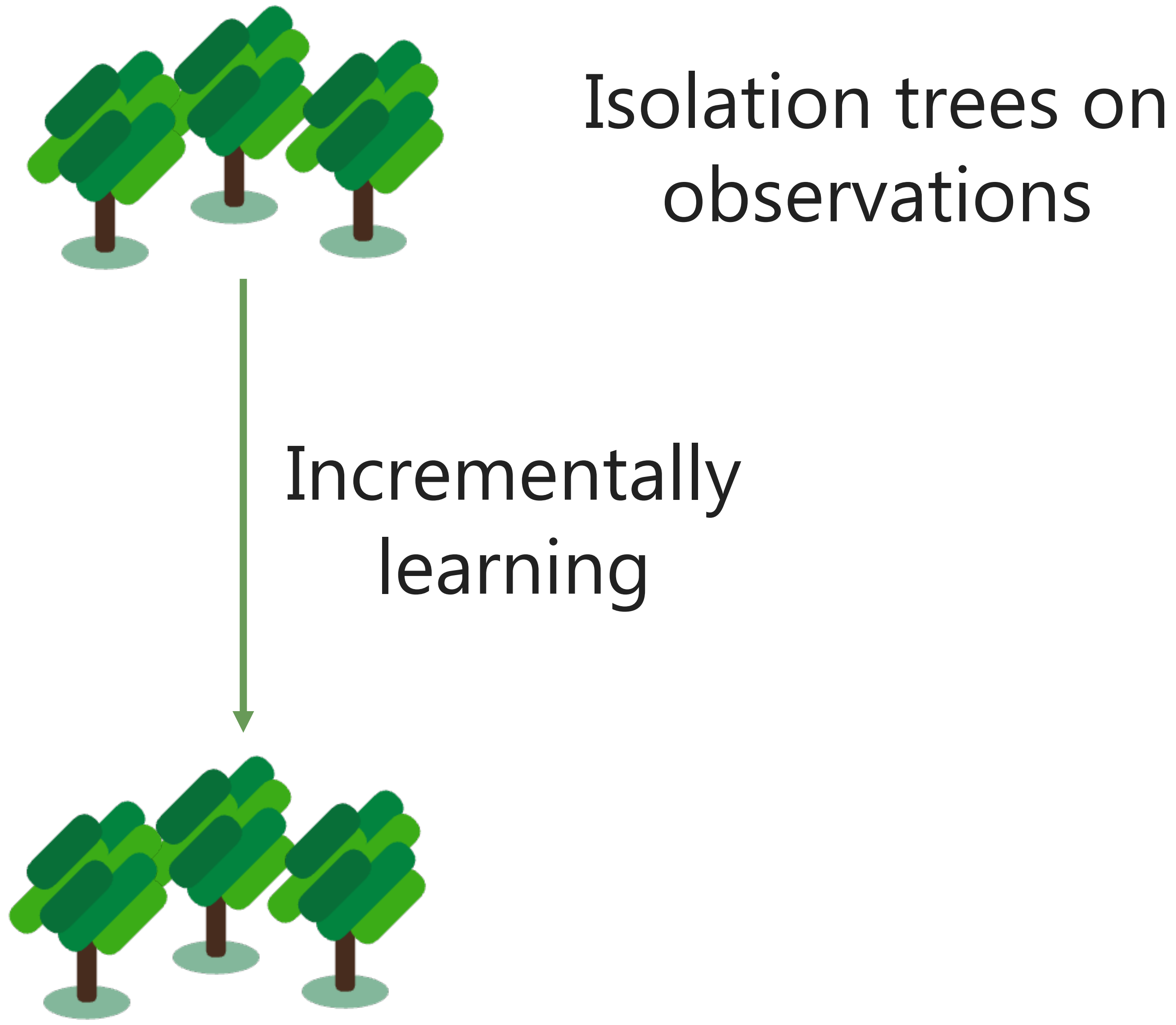
2 Detecting



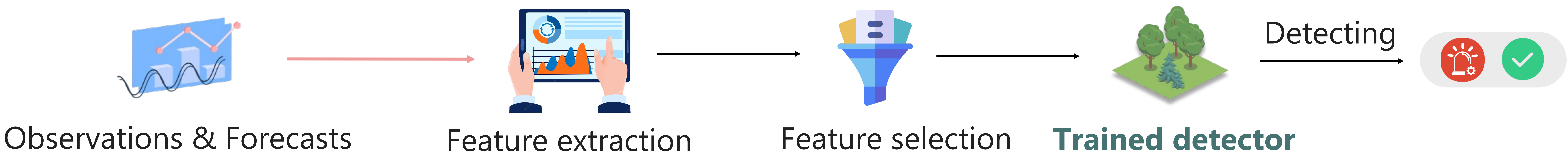
Algorithm 1: Incrementally training isolation forest.

Input: $X^{\text{cat}}_{[1:N]}$, γ, ψ, F_{pre} - previously trained forest
Output: A new forest F consisting of γ trees and F_{pre}

```
Initialize F
i ← 1 while i ≤ γ do
    X′ ← sample(X^cat_[1:N], ψ)
    X′_iso ← F_pre(X′) // Keep the samples "isolated" by F_pre
    F ← F ∪ iTree(X_iso)
end
return F
```



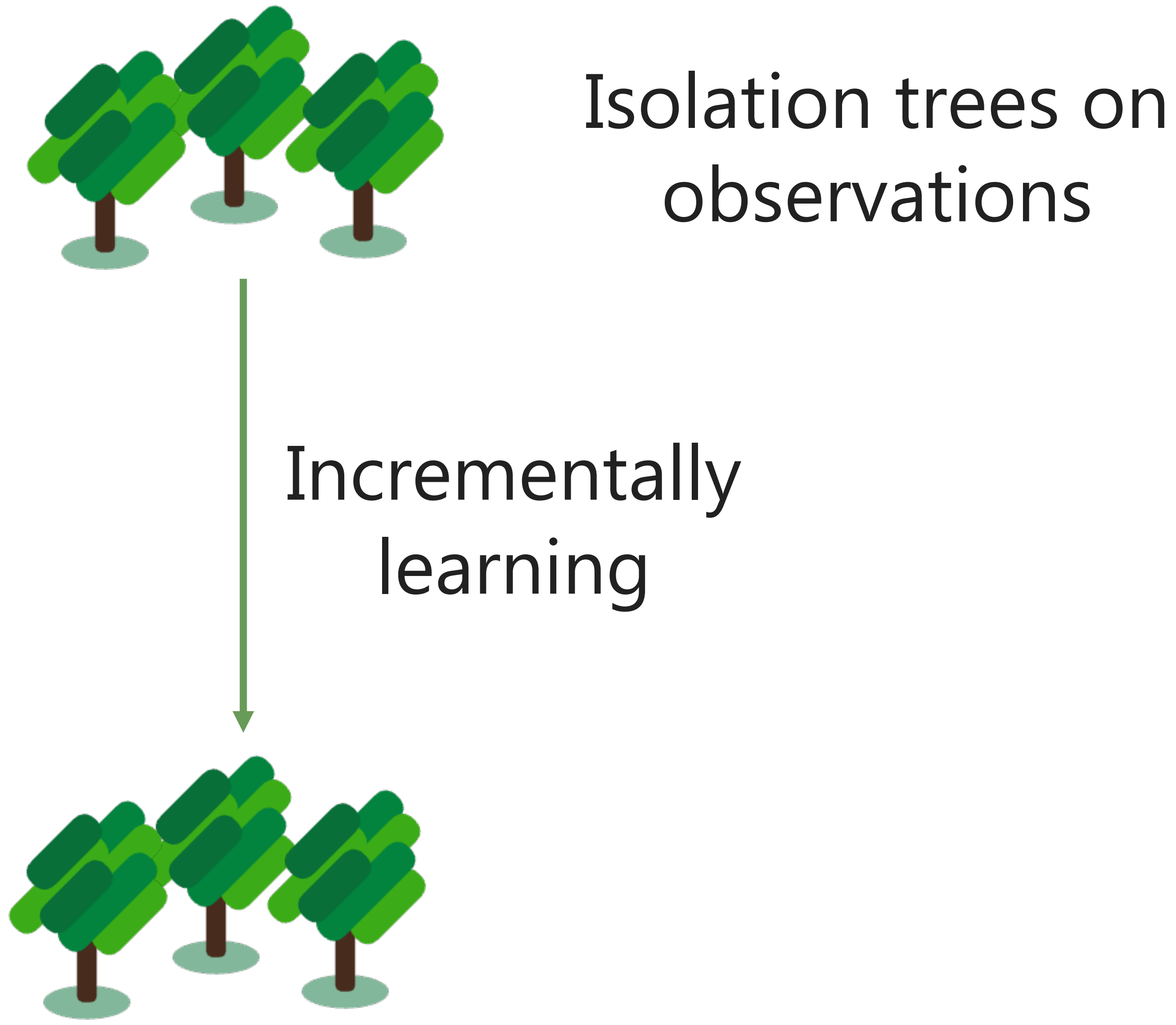
2 Detecting



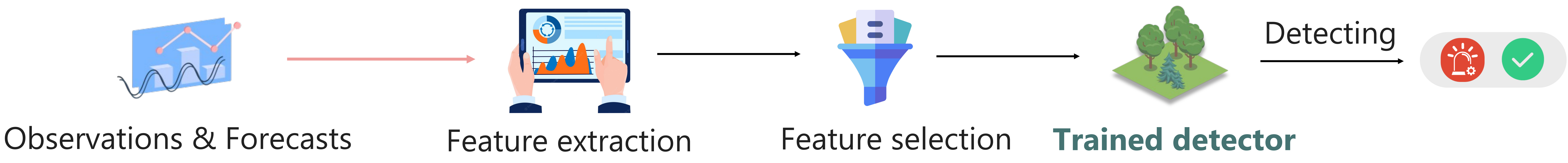
Algorithm 1: Incrementally training isolation forest.

Input: $X^{\mathrm{cat}}_{[1:N]}$, γ, ψ, F_{pre} - previously trained forest
Output: A new forest F consisting of γ trees and F_{pre}

Initialize F
$i \leftarrow 1$ **while** $i \leq \gamma$ **do**
  $X' \leftarrow sample(X^{\mathrm{cat}}_{[1:N]}, \psi)$
  $X'_{iso} \leftarrow F_{pre}(X')$ // Keep the samples "isolated" by F_{pre}
  $F \leftarrow F \cup iTree(X_{iso})$
end
return F



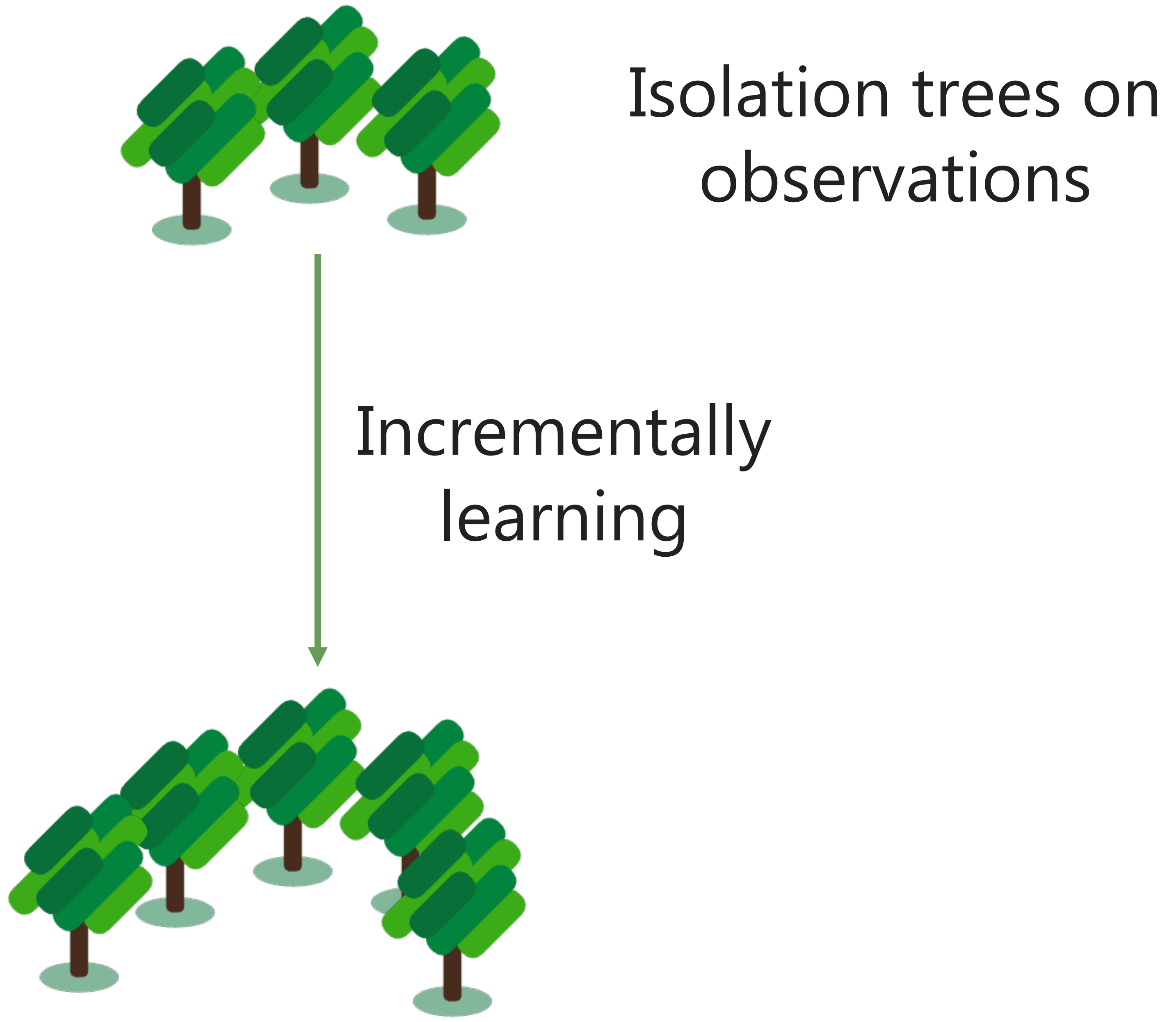
Detecting



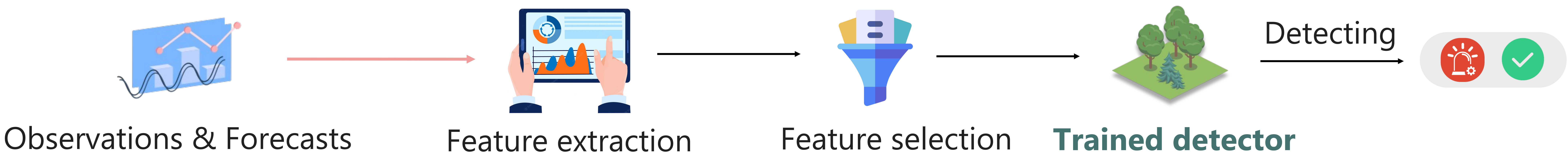
Algorithm 1: Incrementally training isolation forest.

Input: $X^{cat}_{[1:N]}$, γ, ψ, F_{pre} - previously trained forest
Output: A new forest F consisting of γ trees and F_{pre}

Initialize F
$i \leftarrow 1$ **while** $i \leq \gamma$ **do**
  $X' \leftarrow sample(X^{cat}_{[1:N]}, \psi)$
  $X'_{iso} \leftarrow F_{pre}(X')$ // Keep the samples "isolated" by F_{pre}
  $F \leftarrow F \cup iTree(X_{iso})$
end
return F



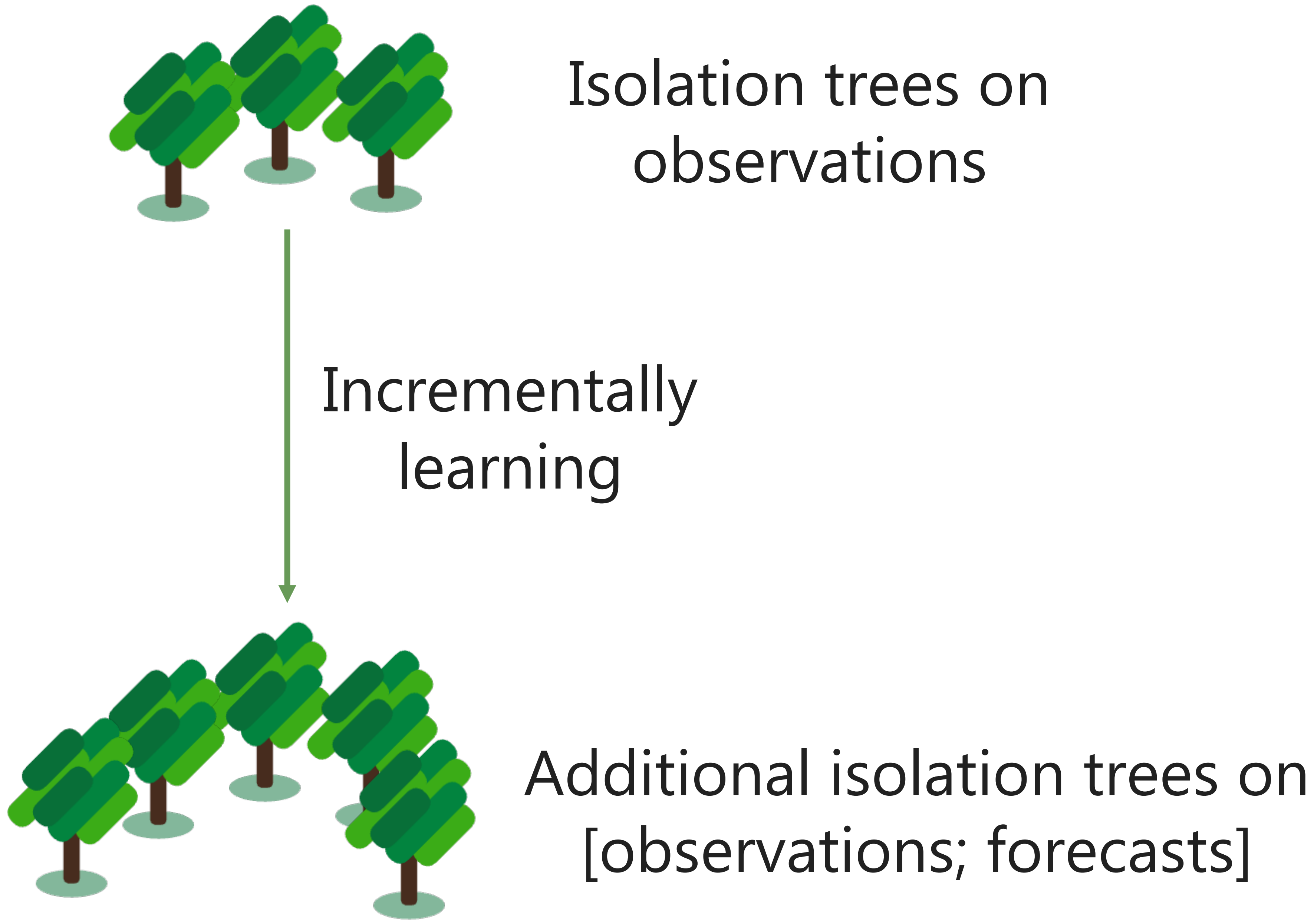
2 Detecting



Algorithm 1: Incrementally training isolation forest.

Input: $X^{\text{cat}}_{[1:N]}$, γ, ψ, F_{pre} - previously trained forest
Output: A new forest F consisting of γ trees and F_{pre}

Initialize F
$i \leftarrow 1$ **while** $i \leq \gamma$ **do**
    $X' \leftarrow sample(X^{\text{cat}}_{[1:N]}, \psi)$
    $X'_{iso} \leftarrow F_{pre}(X')$ // Keep the samples "isolated" by F_{pre}
    $F \leftarrow F \cup iTree(X_{iso})$
end
return F



04 EVALUATION

RQ1: How effective is Maat in anomaly anticipation?

RQ2: How effective is the forecaster of Maat?

RQ3: How much time can Maat advance anomaly alarm?

RQ1: Effectiveness in Anomaly Anticipation

Maat, as a faster-than-real-time anomaly anticipator relying on forecasts,
performs as well as or better than SOTA real-time detectors based on real observations.

OVERALL PERFORMANCE COMPARISON (%)*.

Mode	Methods	AIOps18†			Hades			Yahoo! S5			**Average**		
		F1	*Rec*	*Pre*	*F1*	*Rec*	*Pre*	*F1*	*Rec*	*Pre*	*F1*	*Rec*	*Pre*
Real-time	Dount	36.60	43.06	31.82	49.17	47.49	50.97	58.30	65.77	52.36	48.02	52.11	45.05
	SR-CNN	44.81	**71.91**	32.54	34.25	61.43	23.74	41.06	61.81	30.74	40.04	65.05	29.01
	Adsketch	<u>64.82</u>	64.28	65.37	65.35	57.47	75.73	58.08	67.28	51.09	62.75	63.01	64.06
	Telemanom	49.49	60.10	42.06	46.75	66.29	36.10	54.10	**77.43**	41.57	50.11	67.94	39.91
	LSTM-VAE	46.35	54.57	40.29	36.89	69.07	25.16	62.77	63.35	62.20	48.67	62.33	42.55
	MTAD-GAT	37.85	46.24	32.04	56.90	55.40	58.48	35.62	31.86	40.38	43.46	44.50	43.63
	DAGMM	53.52	58.08	49.63	62.10	55.62	70.29	57.33	51.70	64.33	57.65	55.13	61.42
	OmniAnomaly	57.40	66.82	50.31	68.17	78.81	60.06	53.13	76.75	40.63	59.57	74.13	50.33
	Maat-rt	**66.75**	64.12	**69.60**	**85.30**	84.35	**86.28**	**72.28**	74.65	70.06	**74.78**	**74.37**	**75.31**
FTRT	**Maat**	63.78	58.94	69.48	82.07	**88.77**	76.31	70.31	69.15	**71.51**	72.05	72.29	72.43

RQ2: Effectiveness in Foresting under Anomalies

Maat's forecaster performs effectively in anomalous metric forecasting, reducing MSE by 44.73%~89.81% and sMAPE by 30.76%~65.87% on average.

COMPARISON FOR PERFORMANCE METRIC FORECASTING.

Methods	AIOps18		Hades		Yahoo!S5	
	MSE	*sMAPE*	*MSE*	*sMAPE*	*MSE*	*sMAPE*
GRU	6.170	1.256	3.368	1.957	1.422	1.448
Transformer	5.627	1.400	5.628	1.492	1.717	1.443
TCN	4.610	1.230	3.622	0.835	1.111	1.498
DeepVAR	0.428	0.677	1.250	0.692	0.714	1.022
GRU-MAF	2.607	1.451	6.739	1.959	1.180	1.439
Transformer-MAF	3.235	1.470	2.091	1.677	1.226	1.505
Maat-$\mathcal{F}$	**0.298**	**0.566**	**0.597**	**0.487**	**0.426**	**0.602**

RQ3: Advanced Time of Maat

Maat can anticipate anomalies minutes or hours in advance,
whereas imposing only a few seconds' computation overhead.

The number of advanced sampling intervals

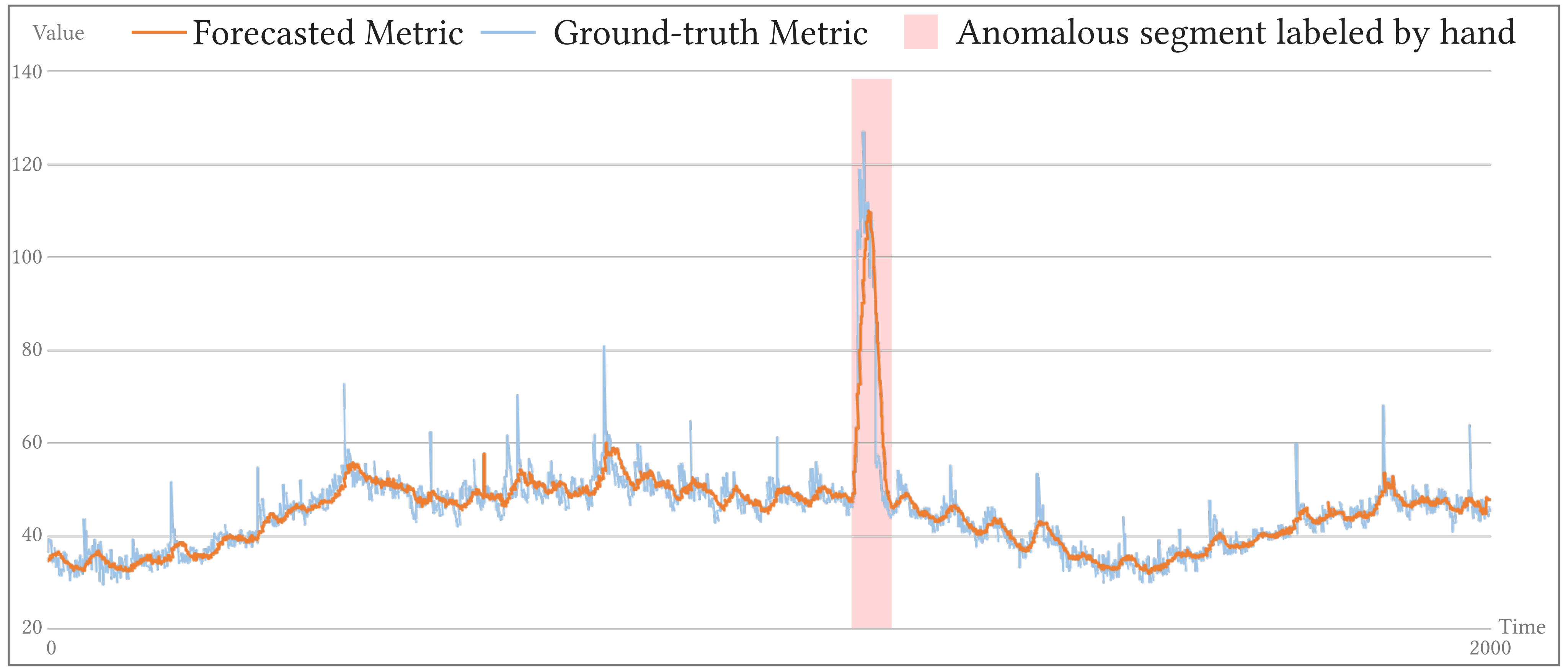
The total overhead / sec

TIME CONSUMPTION OF MAAT (UNIT: SECOND).

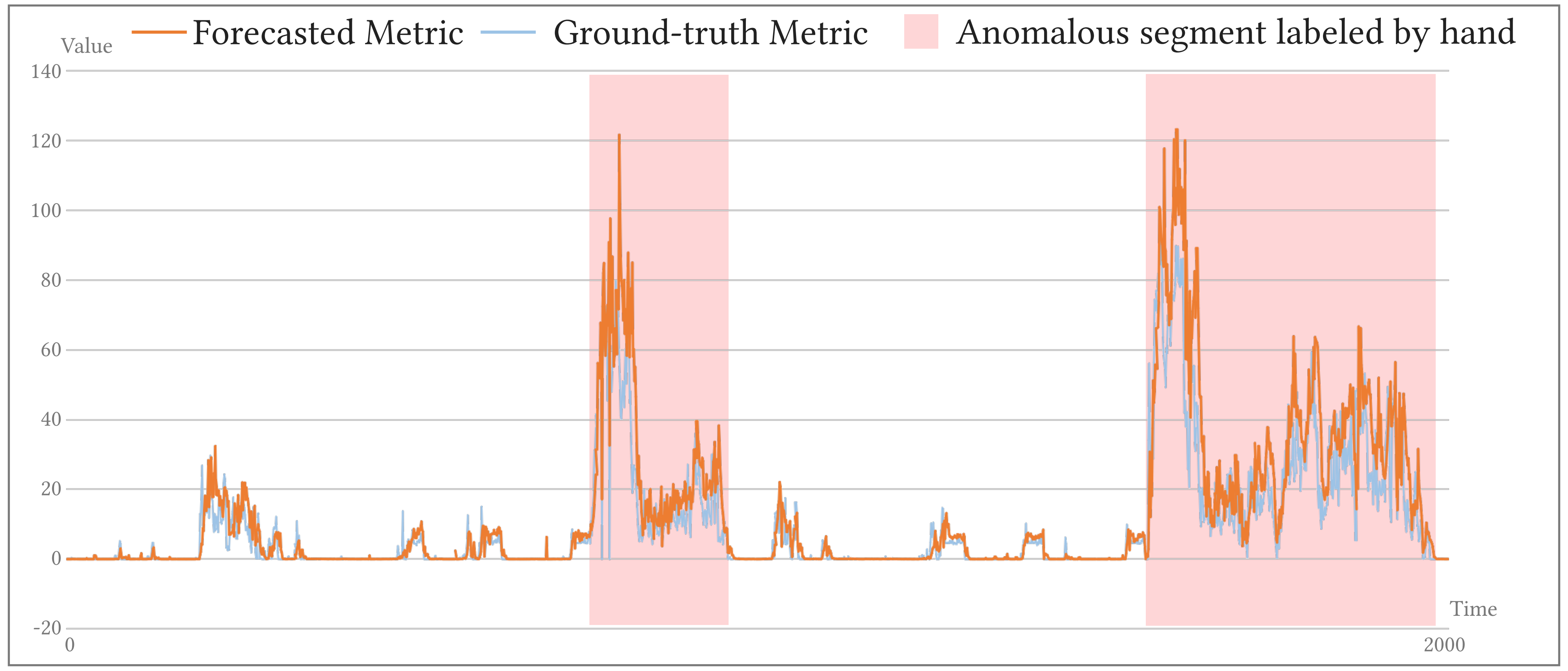
Dataset	#ALen	#PredT	#FeatT	#DeteT	Total
AIOps18	5	3.031	1.320	0.035	4.386
Hades	3	1.922	0.976	0.036	2.934
Yahoo!S5	3	1.915	0.238	0.036	2.189

Successful Cases

Cases of forecasting metrics with anomalies

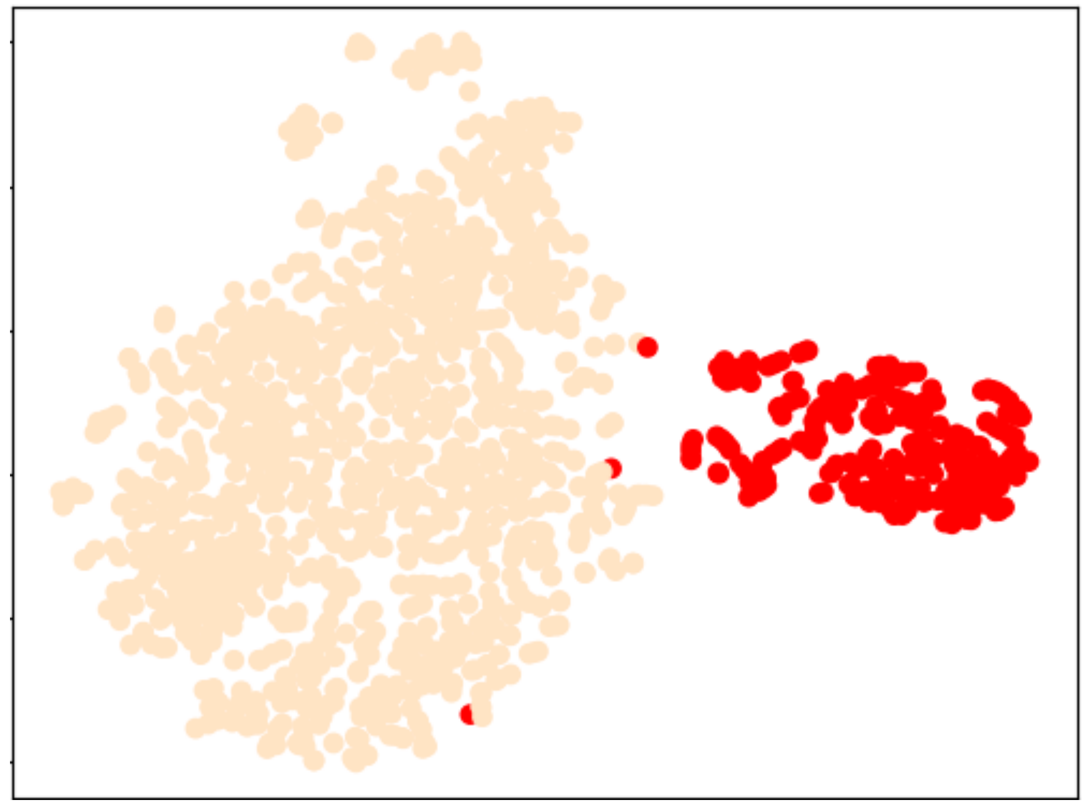


(a) AIOps18: Metric "8723f0fb-eaef-32e6-b372-6034c9c04b80"

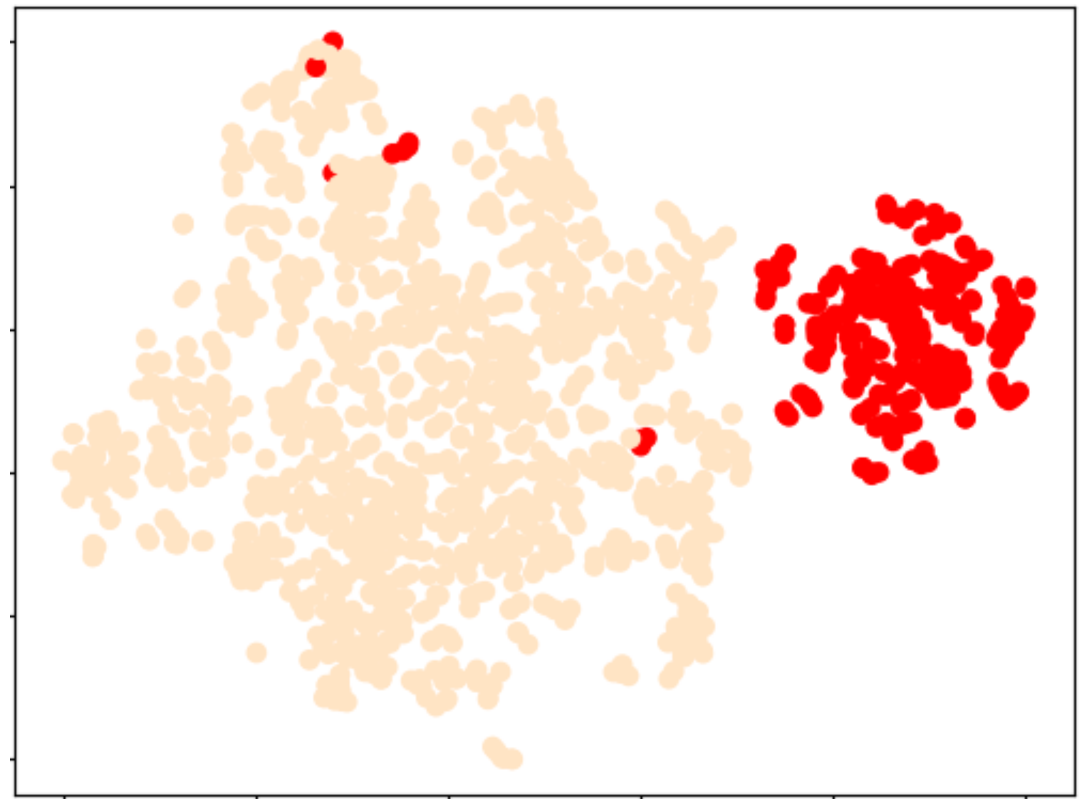


(b) Hades: Metric "CPU iowait"

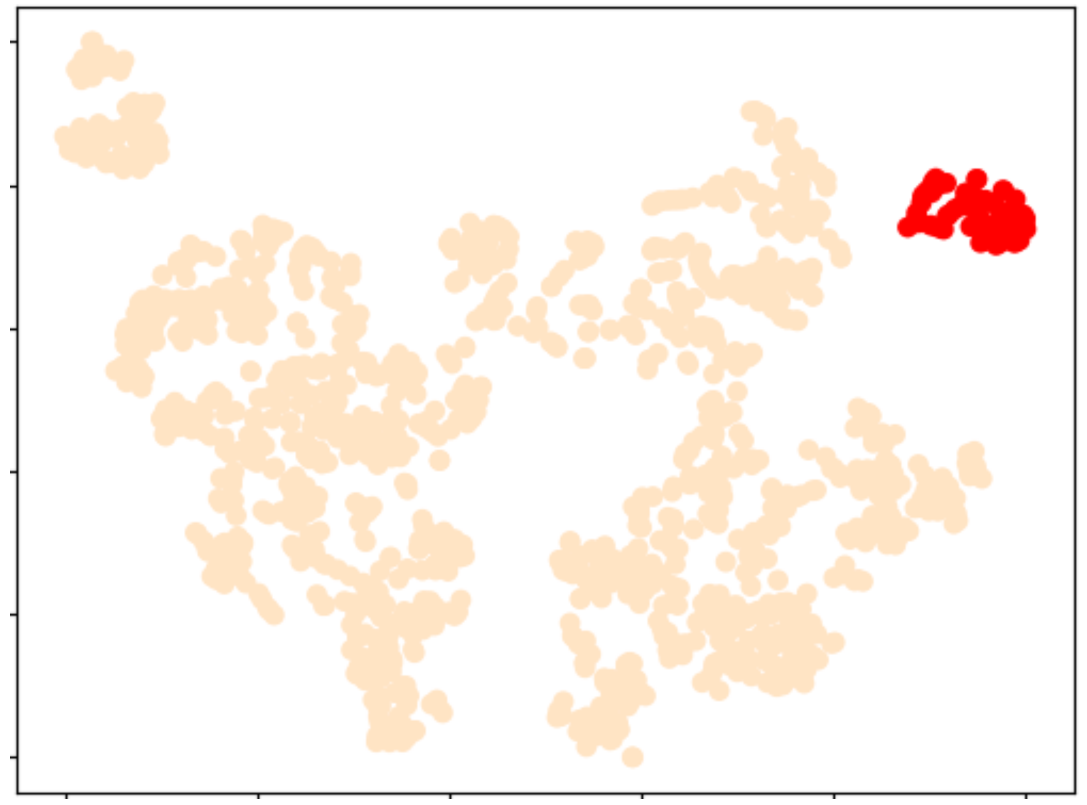
Cases of distinguishing anomalies on features



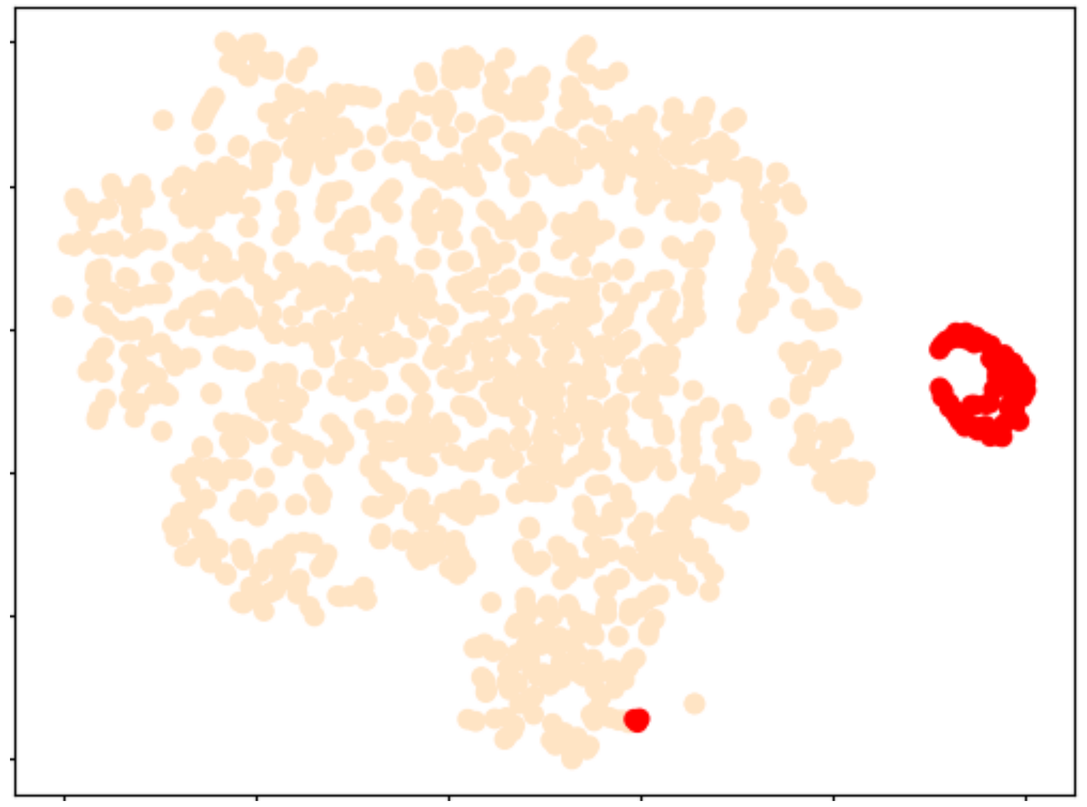
(a) Real-17



(b) Real-19



(c) Real-22



(d) Real-42

ARISE
Automated Reliable Intelligent Software Engineering

香港中文大學
CUHK

THANKS

Presenter: Cheryl LEE

Arise Lab

Full Paper

My Homepage