

A Computational Resources

Specification \ Server	Server 1	Server 2	Server 3
#(CPU)	48	48	48
#(GPU)	4	8	0
CPU Model	Intel(R) Xeon(R) Silver 4214R	Intel(R) Xeon(R) Silver 4116	Intel(R) Xeon(R) Silver 4116
GPU Model	GeForce GTX 1080 Ti	GeForce GTX 1080 Ti	-
Memory Size	92 GB	188GB	64GB

Table 3: Summary of computational resources.

B Datasets

B.1 Synthetic Datasets

The parameters for generating univariate datasets are provided in Table 4. We generate the dataset based on our developed synthetic data generator. Specifically, we generate 20 univariate datasets, where each dataset includes only one type of outlier with different outlier ratio. The outlier ratio is defined as $\frac{\text{labeled_timepoint}}{\text{total}}$. For multivariate, we generate 15 five-dimensional data with 1 to 5 types of outlier in each dimension. We open-source our synthetic dataset in the "benchmark" branch of TODS⁶, where the README.md file gives more details on the datasets. Anyone can reproduce our synthetic datasets by applying above settings.

Type \ Attributes	Length	Dimension	Noise	Outlier Radius	Outlier Ratio	Specifications
Global	200	1	0.05	-	{0.05,0.1,0.15,0.2}	$\lambda=6.0$
Contextual	200	1	0.05	-	{0.05,0.1,0.15,0.2}	$k=5, \eta=2.0$
Shapelet	200	1	0.05	5	{0.05,0.1,0.15,0.2}	$\rho=\text{sgn}(\sin(.))$
Seasonal	200	1	0.05	5	{0.05,0.1,0.15,0.2}	$\omega=4, \hat{\omega}=0.04$
Trend	200	1	0.05	5	{0.05,0.1,0.15,0.2}	$m(\tau)=5, m(\hat{\tau})=0$
Multi-1	200	5	0.05	5	0.05	1 type of outlier
Multi-2	200	5	0.05	5	0.1	2 types of outlier
Multi-3	200	5	0.05	5	0.15	3 types of outlier
Multi-4	200	5	0.05	5	0.20	4 types of outlier
Multi-5	200	5	0.05	5	0.25	5 types of outlier

Table 4: Details of univariate synthetic datasets.

B.2 Real-world Datasets

We involve four multivariate real-world datasets that published within 3 years. Compared to existing multivariate benchmark datasets such as SMD, SMAP, MSL, which only have point outliers, our datasets involve both point and pattern-wise outliers.

- **Credit Card**⁷ is collected by openML which contains transactions made by credit cards in September 2013 by European cardholders. The fraudulent transactions are labeled as outliers. In our repository, we name the processed data as "creditcard". (**License: N/A**)
- **CICIDS**⁸ is collected by Canadian Institute for Cybersecurity in 2017. We adopt the "Thursday-WorkingHours-Morning-WebAttacks" file in 2017 datasets which contains 3 kinds of intrusion attack: XSS, SQL injection and brute force attack. We drop all of the row and columns with NaN and label all kinds of attack as outliers. In our repository, we name the processed data as "web_attack". (**License: CC-BY 4.0**)

⁶<https://github.com/datamllab/tods/tree/benchmark>

⁷<https://www.openml.org/d/1597>

⁸<https://www.unb.ca/cic/datasets/ids-2017.html>

- **GECCO**^[9] is collected by SPOTSeven Lab^[10] for hosting a data challenge^[11] in 2018. We binarized all of the binary columns with 1 and 0. In our repository, we name the processed data as "water_quality". (**License: CC-BY 4.0**)
- **SWAN-SF**^[12] is collected by Harvard Dataverse. We adopt the labeled version from the demo of an open sourced multivariate time series preprocessing toolkit^[13] which can be directly downloaded from bitbucket^[14]. We merge all of labeled csv file as a multivariate time series data. Then, we drop all columns and rows with NaN values, binarize all binary columns and label all of the flares as outliers. In our repository, we name the processed data as "swan_sf". (**License: CC-0**)

The dataset statistics are provided in the table below:

Dataset Attributes	#(Timestamps)	#(Dimensions)	Outlier Ratio
Credit Card	284,807	29	0.173%
Web Attack	170,231	79	1.281%
Water Quality	138,521	10	1.246%
Space Weather	120,000	39	23.8%

Table 5: Details of real-world datasets.

C Hyperparameters

For all of the algorithms, we adopt the implementations from open-sourced outlier detection library: TODS^[15]. Specifically, algorithms developed in PyOD^[16] including GAN, Autoencoder, IForest and OCSVM are adopted into TODS to generate reproducible pipelines. Other adopted algorithms including subsequence segmentation, LSTM-RNN, autoregression and matrix profile are implemented in TODS. For synthetic datasets, we align the contamination ratio for all of the algorithms with the outlier ratio. For real-world datasets, we search the contamination ratios from {0.01, 0.05, 0.1, 0.15, 0.2, 0.25} for all of the algorithms and report the best one. We provide the detailed hyperparameter settings for each algorithm as follows:

- **AR:** We set step=1 and search the window size within {3, 5, 10} for all of the datasets and report the best result.
- **GBRT:** We set step=1 and search the window size within {3, 5, 10} for all of the datasets and report the best result.
- **LSTM-RNN:** For LSTM-RNN, we search the number of hidden layer within {2, 5, 10} and number of hidden units within {32, 64}. Then we train the model for 20 epochs with batch size as 32 and dropout ratio 0.1. Finally, we report the best detection result for each dataset.
- **IForest** We adopt default hyperparameters for IForest, which can be found in TODS^[17]
- **OCSVM** We only set the max_iter=1000 to prevent from infinity iteration. For other hyperparameters, we adopt default hyperparameters for OCSVM which can be found in TODS^[18]
- **Autoencoder:** We search the autoencoder with the candidate neural architectures given by (32,16,32), (16,8,16), (32,16,8,16,32), where the first and last element in each tuple indicate the number of units in the input and output layers, respectively, and other elements indicate that of hidden layers. The autoencoder is trained for 50 epochs with batch size as 32 and using Sigmoid activation function with dropout rate as 0.2.

⁹<https://bit.ly/3f0eRvI>

¹⁰<https://www.spotseven.de>

¹¹<https://www.spotseven.de/gecco/gecco-challenge/gecco-challenge-2018/>

¹²<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/EBCFKM>

¹³https://github.com/AzimAhmadzadeh/mvtsdata_toolkit/blob/master/demo.ipynb

¹⁴https://bitbucket.org/gsudmlab/mvtsdata_toolkit/downloads/petdataset_01.zip

¹⁵[urlhttps://github.com/datamllab/tods/tree/dev](https://github.com/datamllab/tods/tree/dev)

¹⁶[urlhttps://github.com/yzhao062/pyod/](https://github.com/yzhao062/pyod/)

¹⁷https://github.com/datamllab/tods/blob/dev/tods/detection_algorithm/

¹⁸[PyodIsolationForest.py](#)

¹⁸https://github.com/datamllab/tods/blob/dev/tods/detection_algorithm/PyodOCSVM.py

- **GAN:** We adopt the MO-GAAL from TODS which is initially implemented in PyOD, and we tune the hyperparameter k from 1 to 10 and the training epochs within {20, 30, 50}. Since it fails to identify any outliers in real-world datasets, we only report this algorithm on synthetic datasets.
- **Matrix Profile:** We search the window size within {3, 5, 10}, and report the best result with window size as 10. Besides, we only report the results for synthetic datasets because MP cannot complete the training procedure within 1000 CPU hours for all of the real-world datasets.
- **Subsequence Clustering:** We search the window size within {3,5,10} and report the best result with window size as 10 to extract subsequences for establishing \blacktriangle OCSVM and \blacktriangle IForest.

More details with all established pipeline description files as well as experiment results can be found in the "benchmark" branch of TODS ¹⁹

D Detailed Experiment Results

In this section, we provide extensive results on both synthetic and real-world datasets. Specifically, for multivariate synthetic data, since multiple outliers are involved in each dataset, we use **G**, **C**, **S**, **E**, **T** to represent **G**lobal, **C**ontextual, **S**hapelet, **E**asonal and **T**rend outliers, respectively.

Ratio (Global) Metrics	5%			10%			15%			20%		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	1.00	1.00	1.00	0.89	0.96	0.92	0.86	0.89	0.88	0.84	0.86	0.85
GBRT	0.80	0.80	0.80	0.80	0.89	0.92	0.84	0.70	0.75	0.65	0.70	0.68
LSTM-RNN	0.40	0.40	0.40	0.25	0.27	0.26	0.23	0.25	0.24	0.25	0.27	0.26
IForest	0.90	0.90	0.90	0.85	0.89	0.87	0.86	0.90	0.88	0.90	0.92	0.91
OCSVM	0.90	0.90	0.90	0.94	0.89	0.92	0.96	0.89	0.93	0.97	0.92	0.95
AutoEncoder	0.90	0.90	0.90	0.80	0.84	0.82	0.80	0.83	0.81	0.75	0.85	0.76
GAN	0.60	0.60	0.60	0.55	0.61	0.58	0.50	0.54	0.52	0.45	0.49	0.47
\blacktriangle IForest	0.10	0.10	0.10	0.25	0.28	0.26	0.27	0.29	0.28	0.25	0.27	0.26
\blacktriangle OCSVM	0.10	0.10	0.10	0.20	0.22	0.21	0.30	0.32	0.31	0.28	0.30	0.29
MatrixProile	0.20	0.20	0.20	0.05	0.08	0.06	0.16	0.18	0.17	0.23	0.24	0.23

Table 6: Detail experiment results on synthetic univariate time series data with global outliers.

Ratio (Context) Metrics	5%			10%			15%			20%		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.70	0.70	0.70	0.63	0.63	0.63	0.65	0.68	0.67	0.71	0.71	0.71
GBRT	0.50	0.50	0.50	0.50	0.53	0.51	0.53	0.57	0.55	0.58	0.61	0.59
LSTM-RNN	0.00	0.00	0.00	0.05	0.05	0.05	0.13	0.14	0.14	0.15	0.16	0.15
IForest	0.20	0.20	0.20	0.25	0.26	0.26	0.23	0.25	0.24	0.25	0.26	0.27
OCSVM	0.20	0.20	0.20	0.25	0.26	0.26	0.26	0.29	0.28	0.40	0.42	0.41
AutoEncoder	0.00	0.00	0.00	0.05	0.05	0.05	0.06	0.07	0.07	0.23	0.24	0.23
GAN	0.00	0.00	0.00	0.05	0.05	0.05	0.06	0.07	0.07	0.23	0.24	0.23
\blacktriangle IForest	0.10	0.10	0.10	0.20	0.21	0.21	0.27	0.29	0.28	0.33	0.34	0.33
\blacktriangle OCSVM	0.10	0.10	0.10	0.20	0.21	0.21	0.27	0.29	0.28	0.33	0.34	0.33
MatrixProile	0.10	0.10	0.10	0.05	0.05	0.05	0.20	0.21	0.21	0.25	0.26	0.26

Table 7: Detail experiment results on synthetic univariate time series data with contextual outliers.

¹⁹<https://github.com/datamllab/tods/tree/benchmark>

Ratio (Shapelet) Metrics	5%			10%			15%			20%		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.40	0.40	0.40	0.58	0.55	0.56	0.58	0.71	0.64	0.61	0.67	0.64
GBRT	0.90	0.90	0.90	0.60	0.60	0.60	0.57	0.71	0.63	0.53	0.62	0.57
LSTM-RNN	0.00	0.00	0.00	0.15	0.15	0.15	0.10	0.13	0.11	0.28	0.32	0.30
IForest	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.04	0.03	0.00	0.00	0.00
OCSVM	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.08	0.06	0.08	0.11	0.10
AutoEncoder	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.08	0.07	0.18	0.21	0.19
GAN	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.08	0.07	0.18	0.21	0.19
▲IForest	0.00	0.00	0.00	0.30	0.30	0.30	0.43	0.54	0.48	0.45	0.53	0.49
▲OCSVM	0.00	0.00	0.00	0.50	0.50	0.50	0.50	0.63	0.56	0.56	0.68	0.62
MatrixProile	0.40	0.40	0.40	0.40	0.40	0.40	0.57	0.71	0.63	0.58	0.68	0.62

Table 8: Detail experiment results on synthetic univariate time series data with shapelet outliers.

▲ represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Ratio (Seasonal) Metrics	5%			10%			15%			20%		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.50	0.50	0.50	0.42	0.40	0.41	0.48	0.58	0.53	0.53	0.59	0.56
GBRT	0.40	0.40	0.40	0.50	0.50	0.50	0.46	0.58	0.52	0.55	0.65	0.60
LSTM-RNN	0.40	0.40	0.40	0.55	0.55	0.55	0.56	0.71	0.63	0.55	0.65	0.59
IForest	0.05	0.10	0.07	0.13	0.20	0.16	0.05	0.08	0.06	0.12	0.18	0.14
OCSVM	0.05	0.10	0.07	0.10	0.15	0.12	0.13	0.21	0.15	0.18	0.26	0.21
AutoEncoder	0.00	0.00	0.00	0.10	0.10	0.10	0.16	0.21	0.19	0.18	0.21	0.19
GAN	0.00	0.00	0.00	0.10	0.10	0.10	0.17	0.21	0.19	0.18	0.21	0.19
▲IForest	0.60	0.60	0.60	0.55	0.55	0.55	0.60	0.75	0.66	0.53	0.62	0.57
▲OCSVM	0.60	0.60	0.60	0.55	0.55	0.55	0.60	0.75	0.66	0.55	0.65	0.60
MatrixProile	0.60	0.60	0.60	0.45	0.45	0.45	0.60	0.75	0.67	0.53	0.62	0.57

Table 9: Detail experiment results on synthetic univariate time series data with seasonal outliers.

▲ represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Ratio (Trend) Metrics	5%			10%			15%			20%		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.20	0.20	0.20	0.42	0.40	0.41	0.31	0.38	0.34	0.37	0.41	0.39
GBRT	0.50	0.50	0.50	0.55	0.55	0.55	0.66	0.83	0.74	0.60	0.71	0.65
LSTM-RNN	0.00	0.00	0.00	0.15	0.15	0.15	0.30	0.29	0.29	0.58	0.68	0.62
IForest	0.20	0.20	0.20	0.15	0.15	0.15	0.23	0.29	0.26	0.10	0.12	0.11
OCSVM	0.05	0.10	0.07	0.10	0.15	0.12	0.13	0.21	0.15	0.18	0.26	0.21
AutoEncoder	0.20	0.20	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GAN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
▲IForest	0.00	0.00	0.00	0.55	0.55	0.55	0.47	0.58	0.52	0.38	0.44	0.41
▲OCSVM	0.00	0.00	0.00	0.85	0.85	0.85	0.53	0.67	0.59	0.50	0.59	0.54
MatrixProile	0.20	0.20	0.20	0.00	0.00	0.00	0.43	0.54	0.48	0.55	0.64	0.59

Table 10: Detail experiment results on synthetic univariate time series data with trend outliers.

▲ represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Outlier Types (Single-2) Metrics	GC			CS			SE			ET		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.74	0.74	0.74	0.69	0.68	0.68	0.51	0.50	0.51	0.63	0.63	0.63
GBRT	0.70	0.72	0.71	0.58	0.58	0.58	0.70	0.70	0.70	0.43	0.43	0.43
LSTM-RNN	0.18	0.18	0.18	0.21	0.56	0.21	0.43	0.43	0.43	0.43	0.43	0.43
IForest	0.66	0.67	0.66	0.05	0.05	0.05	0.10	0.10	0.10	0.28	0.28	0.28
OCSVM	0.70	0.72	0.71	0.08	0.08	0.08	0.08	0.08	0.08	0.30	0.30	0.30
AutoEncoder	0.48	0.49	0.48	0.23	0.23	0.23	0.05	0.05	0.05	0.30	0.30	0.30
GAN	0.40	0.41	0.41	0.03	0.03	0.03	0.05	0.05	0.05	0.05	0.05	0.05
▲IForest	0.15	0.15	0.15	0.48	0.48	0.48	0.68	0.68	0.68	0.28	0.28	0.28
▲OCSVM	0.15	0.15	0.15	0.58	0.58	0.58	0.70	0.70	0.70	0.30	0.30	0.30
MatrixProile	0.18	0.18	0.18	0.20	0.20	0.20	0.58	0.58	0.58	0.40	0.40	0.40

Table 11: Detail experiment results on univariate time series data with two type of outliers. **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively; **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively and **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Outlier Types (Single-3) Metrics	GCS			CSE			SET		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.63	0.66	0.64	0.53	0.57	0.55	0.55	0.55	0.55
GBRT	0.60	0.64	0.62	0.48	0.52	0.50	0.70	0.70	0.70
LSTM-RNN	0.19	0.20	0.20	0.36	0.34	0.35	0.23	0.23	0.23
IForest	0.33	0.36	0.35	0.15	0.16	0.16	0.35	0.35	0.35
OCSVM	0.47	0.50	0.48	0.17	0.18	0.17	0.30	0.30	0.30
AutoEncoder	0.35	0.38	0.36	0.08	0.09	0.09	0.18	0.18	0.18
GAN	0.27	0.29	0.28	0.08	0.09	0.09	0.26	0.26	0.26
▲IForest	0.17	0.18	0.17	0.55	0.59	0.57	0.35	0.35	0.35
▲OCSVM	0.27	0.29	0.28	0.57	0.61	0.59	0.33	0.33	0.33
MatrixProile	0.27	0.29	0.28	0.42	0.45	0.43	0.53	0.53	0.53

Table 12: Detail experiment results on univariate time series data with three types of outliers. **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively and **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Outlier Types (Single-4&5) Metrics	GCSE			CSET			GCSET		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.60	0.66	0.63	0.54	0.57	0.55	0.61	0.79	0.65
GBRT	0.54	0.61	0.57	0.53	0.55	0.54	0.59	0.66	0.62
LSTM-RNN	0.30	0.34	0.32	0.26	0.28	0.27	0.34	0.38	0.36
IForest	0.40	0.45	0.42	0.34	0.36	0.35	0.45	0.50	0.47
OCSVM	0.36	0.41	0.38	0.35	0.37	0.36	0.44	0.49	0.46
AutoEncoder	0.31	0.35	0.33	0.18	0.18	0.18	0.32	0.36	0.34
GAN	0.21	0.24	0.23	0.20	0.21	0.21	0.37	0.41	0.39
▲IForest	0.34	0.38	0.36	0.35	0.37	0.36	0.33	0.37	0.35
▲OCSVM	0.45	0.51	0.48	0.35	0.37	0.36	0.27	0.30	0.28
MatrixProile	0.34	0.38	0.36	0.48	0.50	0.49	0.41	0.46	0.43

Table 13: Detail experiment results on univariate time series data with four and all types of outliers. **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively; and **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Outlier Type (Multi-1) Metrics	Global			Contextual			Shapelet			SEasonal			Trend		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.45	0.45	0.45	0.45	0.45	0.45	0.30	0.30	0.30	0.30	0.30	0.30	0.10	0.10	0.10
GBRT	0.40	0.40	0.40	0.35	0.35	0.35	0.25	0.25	0.25	0.35	0.35	0.35	0.25	0.25	0.25
LSTM-RNN	0.15	0.15	0.15	0.20	0.20	0.20	0.57	0.40	0.47	0.55	0.50	0.53	0.64	0.70	0.67
IForest	0.85	0.85	0.85	0.60	0.60	0.60	0.35	0.35	0.35	0.70	0.70	0.70	0.30	0.30	0.30
OCSVM	0.90	0.90	0.90	0.75	0.75	0.75	0.35	0.35	0.35	0.70	0.70	0.70	0.45	0.45	0.45
AutoEncoder	0.20	0.20	0.20	0.05	0.05	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GAN	0.19	0.15	0.17	0.15	0.15	0.15	0.00	0.00	0.00	0.25	0.25	0.25	0.00	0.00	0.00
▲IForest	0.10	0.10	0.10	0.10	0.10	0.10	0.25	0.25	0.25	0.35	0.35	0.35	0.35	0.35	0.35
▲OCSVM	0.10	0.10	0.10	0.10	0.10	0.10	0.20	0.20	0.20	0.45	0.45	0.45	0.30	0.30	0.30
MatrixProile	0.10	0.10	0.10	0.05	0.05	0.05	0.05	0.05	0.05	0.55	0.55	0.55	0.20	0.20	0.20

Table 14: Detail experiment results on multivariate time series data with only one type of outliers. **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively and **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Outlier Types (Multi-2) Metrics	GC			CS			SE			ET		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.62	0.62	0.62	0.41	0.42	0.42	0.31	0.30	0.30	0.41	0.43	0.42
GBRT	0.55	0.56	0.56	0.25	0.26	0.26	0.28	0.28	0.28	0.35	0.38	0.36
LSTM-RNN	0.15	0.15	0.15	0.13	0.11	0.12	0.53	0.45	0.49	0.67	0.70	0.68
IForest	0.75	0.77	0.76	0.53	0.55	0.54	0.45	0.45	0.45	0.73	0.78	0.75
OCSVM	0.80	0.82	0.81	0.43	0.45	0.44	0.53	0.53	0.53	0.73	0.78	0.75
AutoEncoder	0.20	0.21	0.20	0.05	0.05	0.05	0.08	0.08	0.08	0.10	0.11	0.10
GAN	0.00	0.00	0.00	0.10	0.11	0.10	0.08	0.08	0.08	0.13	0.14	0.13
▲IForest	0.15	0.15	0.15	0.25	0.26	0.26	0.43	0.43	0.43	0.53	0.57	0.55
▲OCSVM	0.25	0.26	0.25	0.20	0.21	0.21	0.33	0.33	0.33	0.60	0.65	0.62
MatrixProile	0.28	0.28	0.28	0.13	0.13	0.13	0.13	0.13	0.13	0.41	0.38	0.39

Table 15: Detail experiment results on multivariate time series data with only two type of outliers. **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively; **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively and **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Outlier Types (Multi-3) Metrics	GCS			CSE			SET		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.47	0.49	0.48	0.37	0.38	0.37	0.35	0.35	0.35
GBRT	0.45	0.47	0.46	0.43	0.45	0.44	0.32	0.32	0.32
LSTM-RNN	0.13	0.13	0.13	0.16	0.16	0.16	0.27	0.27	0.27
IForest	0.60	0.63	0.62	0.58	0.60	0.59	0.38	0.38	0.38
OCSVM	0.63	0.67	0.65	0.53	0.55	0.54	0.43	0.43	0.43
AutoEncoder	0.13	0.14	0.14	0.10	0.10	0.10	0.13	0.13	0.13
GAN	0.08	0.09	0.09	0.08	0.09	0.09	0.00	0.00	0.00
▲IForest	0.25	0.26	0.26	0.33	0.34	0.34	0.27	0.27	0.27
▲OCSVM	0.23	0.25	0.24	0.40	0.41	0.41	0.38	0.38	0.38
MatrixProile	0.33	0.35	0.34	0.27	0.28	0.27	0.28	0.28	0.28

Table 16: Detail experiment results on multivariate time series data with three types of outliers. **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively and **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Outlier Types (Multi-4&5) Metrics	GCSE			CSET			GCSET		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.54	0.56	0.55	0.53	0.60	0.56	0.59	0.77	0.64
GBRT	0.48	0.51	0.49	0.36	0.43	0.39	0.47	0.56	0.51
LSTM-RNN	0.24	0.24	0.24	0.29	0.32	0.31	0.22	0.26	0.24
IForest	0.63	0.67	0.65	0.63	0.74	0.68	0.48	0.57	0.52
OCSVM	0.68	0.72	0.70	0.55	0.65	0.59	0.62	0.74	0.67
AutoEncoder	0.16	0.17	0.17	0.23	0.26	0.24	0.20	0.24	0.22
GAN	0.16	0.17	0.17	0.18	0.21	0.19	0.15	0.15	0.15
▲IForest	0.40	0.43	0.41	0.53	0.62	0.57	0.34	0.45	0.39
▲OCSVM	0.40	0.43	0.41	0.53	0.62	0.57	0.37	0.49	0.42
MatrixProfile	0.39	0.41	0.40	0.34	0.40	0.36	0.43	0.51	0.47

Table 17: Detail experiment results on multivariate time series data with four and all types of outliers. **G, C, S, E, T** represent global, contextual, shapelet, seasonal and trend outliers respectively; and **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Dataset (0.01) Metrics	Credit Card			CICIDS			GECCO			SWAN-SF		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.113	0.652	0.192	0.000	0.000	0.000	0.392	0.314	0.349	0.831	0.034	0.067
GBRT	0.113	0.657	0.193	0.000	0.000	0.000	0.175	0.140	0.156	0.848	0.036	0.068
LSTM-RNN	0.002	0.016	0.005	0.000	0.000	0.00	0.343	0.275	0.305	0.621	0.026	0.050
IForest	0.098	0.569	0.168	0.009	0.007	0.008	0.439	0.353	0.391	0.975	0.041	0.079
OCSVM	0.107	0.620	0.183	0.000	0.000	0.000	0.481	0.386	0.428	0.799	0.034	0.064
AutoEncoder	0.103	0.598	0.176	0.000	0.000	0.000	0.424	0.340	0.377	0.718	0.030	0.058
▲IForest	0.039	0.226	0.066	0.005	0.004	0.005	0.392	0.315	0.390	0.998	0.041	0.079
▲OCSVM	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.004	0.005	0.193	0.001	0.001
MatrixProfile	0.003	0.022	0.007	0.003	0.002	0.003	0.075	0.061	0.068	0.778	0.033	0.063

Table 18: Benchmark results on four real-world multivariate sequential data. **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Dataset (0.05) Metrics	Credit Card			CICIDS			GECCO			SWAN-SF		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.030	0.860	0.057	0.003	0.012	0.005	0.131	0.527	0.210	0.695	0.146	0.241
GBRT	0.030	0.862	0.058	0.003	0.013	0.005	0.073	0.292	0.117	0.722	0.152	0.251
LSTM-RNN	0.004	0.110	0.007	0.016	0.063	0.026	0.128	0.513	0.205	0.390	0.082	0.135
IForest	0.029	0.854	0.057	0.010	0.040	0.016	0.174	0.698	0.279	0.862	0.181	0.299
OCSVM	0.030	0.878	0.059	0.002	0.009	0.004	0.185	0.743	0.296	0.731	0.154	0.254
AutoEncoder	0.030	0.878	0.059	0.011	0.042	0.017	0.155	0.623	0.248	0.781	0.164	0.271
▲IForest	0.017	0.494	0.033	0.006	0.022	0.009	0.142	0.570	0.227	0.660	0.140	0.230
▲OCSVM	0.000	0.000	0.000	0.000	0.000	0.000	0.33	0.008	0.035	0.193	0.001	0.001
MatrixProfile	0.006	0.185	0.012	0.003	0.011	0.005	0.046	0.185	0.074	0.335	0.071	0.116

Table 19: Benchmark results on four real-world multivariate sequential data. **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Dataset (0.1) Metrics	Credit Card			CICIDS			GECCO			SWAN-SF		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.016	0.898	0.031	0.009	0.067	0.015	0.076	0.604	0.134	0.566	0.238	0.335
GBRT	0.015	0.896	0.030	0.009	0.069	0.016	0.050	0.399	0.088	0.616	0.259	0.364
LSTM-RNN	0.003	0.161	0.005	0.022	0.172	0.039	0.086	0.689	0.153	0.527	0.221	0.312
IForest	0.015	0.880	0.030	0.006	0.045	0.010	0.104	0.769	0.185	0.734	0.308	0.434
OCSVM	0.016	0.904	0.031	0.006	0.044	0.010	0.101	0.809	0.179	0.553	0.232	0.327
AutoEncoder	0.016	0.902	0.031	0.006	0.044	0.010	0.092	0.742	0.164	0.585	0.246	0.346
▲IForest	0.008	0.453	0.015	0.007	0.058	0.013	0.096	0.769	0.175	0.525	0.221	0.311
▲OCSVM	0.002	0.132	0.004	0.000	0.000	0.000	0.020	0.163	0.036	0.193	0.001	0.001
MatrixProfile	0.005	0.305	0.010	0.005	0.041	0.009	0.036	0.286	0.063	0.218	0.092	0.129

Table 20: Benchmark results on four real-world multivariate sequential data. **▲** represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Dataset (0.15) Metrics	Credit Card			CICIDS			GECCO			SWAN-SF		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.010	0.911	0.021	0.011	0.133	0.021	0.054	0.647	0.099	0.484	0.305	0.374
GBRT	0.011	0.915	0.021	0.011	0.131	0.021	0.039	0.474	0.073	0.523	0.330	0.405
LSTM-RNN	0.002	0.197	0.005	0.016	0.190	0.030	0.065	0.776	0.119	0.264	0.166	0.204
IForest	0.010	0.905	0.021	0.004	0.045	0.007	0.073	0.529	0.135	0.655	0.413	0.506
OCSVM	0.011	0.917	0.021	0.004	0.046	0.007	0.071	0.860	0.132	0.544	0.343	0.424
AutoEncoder	0.011	0.911	0.021	0.004	0.046	0.007	0.065	0.784	0.120	0.554	0.349	0.428
▲IForest	0.008	0.681	0.016	0.008	0.094	0.015	0.069	0.833	0.128	0.491	0.309	0.380
▲OCSVM	0.002	0.181	0.004	0.000	0.000	0.000	0.021	0.256	0.039	0.193	0.001	0.001
MatrixProfile	0.005	0.398	0.009	0.007	0.080	0.013	0.031	0.373	0.057	0.186	0.117	0.144

Table 21: Benchmark results on four real-world multivariate sequential data. ▲ represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Dataset (0.2) Metrics	Credit Card			CICIDS			GECCO			SWAN-SF		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.008	0.923	0.016	0.012	0.194	0.023	0.042	0.681	0.080	0.421	0.354	0.385
GBRT	0.008	0.927	0.016	0.014	0.213	0.026	0.033	0.537	0.063	0.447	0.375	0.408
LSTM-RNN	0.003	0.360	0.006	0.024	0.383	0.046	0.046	0.746	0.087	0.305	0.256	0.278
IForest	0.008	0.931	0.159	0.003	0.047	0.006	0.059	0.954	0.112	0.587	0.494	0.537
OCSVM	0.008	0.933	0.016	0.003	0.046	0.006	0.057	0.910	0.107	0.505	0.424	0.461
AutoEncoder	0.008	0.929	0.016	0.003	0.046	0.006	0.053	0.849	0.099	0.535	0.449	0.488
▲IForest	0.007	0.783	0.013	0.011	0.168	0.020	0.051	0.820	0.096	0.425	0.358	0.388
▲OCSVM	0.002	0.240	0.004	0.000	0.000	0.000	0.021	0.341	0.040	0.193	0.001	0.001
MatrixProfile	0.004	0.459	0.008	0.005	0.087	0.010	0.028	0.442	0.052	0.173	0.146	0.158

Table 22: Benchmark results on four real-world multivariate sequential data. ▲ represents that the subsequence segmentation is adopted as the feature input for the algorithm.

Dataset (0.25) Metrics	Credit Card			CICIDS			GECCO			SWAN-SF		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AR	0.006	0.935	0.013	0.016	0.310	0.030	0.036	0.712	0.068	0.375	0.394	0.384
GBRT	0.007	0.935	0.013	0.018	0.351	0.034	0.029	0.586	0.056	0.390	0.409	0.399
LSTM-RNN	0.002	0.337	0.005	0.022	0.427	0.042	0.035	0.694	0.066	0.237	0.249	0.243
IForest	0.007	0.945	0.013	0.004	0.084	0.008	0.048	0.956	0.091	0.569	0.598	0.583
OCSVM	0.007	0.945	0.013	0.003	0.049	0.005	0.047	0.940	0.089	0.474	0.498	0.485
AutoEncoder	0.007	0.943	0.013	0.003	0.049	0.005	0.045	0.910	0.086	0.497	0.522	0.509
▲IForest	0.006	0.797	0.011	0.011	0.221	0.022	0.046	0.931	0.088	0.406	0.425	0.416
▲OCSVM	0.002	0.305	0.004	0.000	0.000	0.000	0.018	0.362	0.034	0.193	0.001	0.001
MatrixProfile	0.004	0.514	0.007	0.006	0.121	0.012	0.025	0.492	0.047	0.167	0.175	0.171

Table 23: Benchmark results on four real-world multivariate sequential data. ▲ represents that the subsequence segmentation is adopted as the feature input for the algorithm.