

TS2Vec : Towards Universal Representation of Time Series

<https://arxiv.org/pdf/2106.10466>

0. Introduction

- 시계열 데이터는 금융, 수요 예측, 기후 모델링 등 다양한 산업에서 핵심 역할을 함.
- 기존 시계열 표현 학습은 보통 인스턴스 수준 표현만 학습하므로, 예측·이상 탐지 같이 세부 시점 정보가 필요한 과제에 한계가 있음.
- 본 논문은 TS2Vec라는 계층적 대비 학습을 통해 시계열의 다양한 의미 수준 표현을 학습할 수 있는 프레임워크를 제안.
- 주요 기여
 - 모든 의미 수준(타임스탬프부터 전체 시리즈)에서의 표현 학습 가능
 - 증가된 견고성 및 다양한 downstream task(분류·예측·이상 탐지)에서 우수한 성능

1. Overview

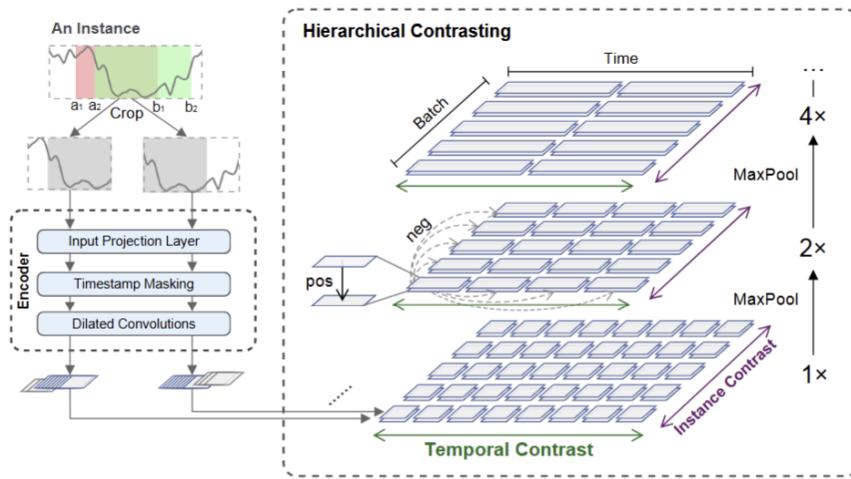
- 서로 다른 확대된 문맥(context) 뷔를 생성하고, 계층적 대비 학습을 통해 임의의 시점 또는 구간에 대한 의미 있는 표현을 얻음
- 서로 다른 샘플 간 비교뿐 아니라 같은 타임스탬프의 두 문맥 표현 간 일관성 학습
- 임의의 하위 시퀀스 표현은 해당 타임스탬프 표현들의 단순 집계(max pooling)로 계산
- 분류, 예측, 이상 탐지 같은 다양한 태스크에서 단일 통일된 표현으로 적용 가능

2. Challenges

- 세부 정보 손실 : 기존 방법은 전체 인스턴스 대표값만 학습하여 세부 상황 파악이 어려움

- 기존 대비 학습 방식 제한 : 단일 수준 대비만 적용하여 분류 성능 중심으로 제한됨
- 시계열 특성 문제 : 이미지나 텍스트 대비에서 사용된 augment 전략이 항상 적합하지 않음

3. Method



- Input Processing & Architecture:

시계열 입력을 타임스탬프별 임베딩으로 변환하고, dilated convolution 기반 인코더를 통해 문맥적 시계열 표현 추출

- Context View Generation:

타임스탬프 마스킹과 랜덤 크롭을 통해 서로 다른 문맥 뷰 생성

- Temporal Contrastive Loss:

서로 다른 뷰에서 동일 타임스탬프의 표현을 양성(positive) 샘플로 사용

- Instance Contrastive Loss:

같은 타임스탬프에서 다른 인스턴스 표현을 음성(negative) 샘플로 사용

- Hierarchical Contrasting:

여러 의미 수준(타임스탬프 → 집계된 시퀀스 → 전체 인스턴스)에서 반복 대비 학습 수행

4. Experiments

- 데이터셋
 - UCR 시계열 아카이브 — 125개 univariate
 - UEA 시계열 아카이브 — 29개 multivariate
- Downstream Tasks
 - 분류 : SVM 분류기 위에 TS2Vec 표현 사용
 - 예측 : 표현 기반 선형 회귀
 - 이상 탐지 : 표현에 기반한 비지도 이상 탐지
- 비교 대상 : T-Loss, TS-TCC, TNC 등 기존 시계열 대비 학습 방법
- 평가 지표 : 정확도, MSE, F1 등 task별 적합한 지표 사용

5. Results

Dataset	H	TS2Vec	Informer	LogTrans	N-BEATS	TCN	LSTnet
ETTh ₁	24	0.039	0.098	0.103	0.094	0.075	0.108
	48	0.062	0.158	0.167	0.210	0.227	0.175
	168	0.134	0.183	0.207	0.232	0.316	0.396
	336	0.154	0.222	0.230	0.232	0.306	0.468
	720	0.163	0.269	0.273	0.322	0.390	0.659
ETTh ₂	24	0.090	0.093	0.102	0.198	0.103	3.554
	48	0.124	0.155	0.169	0.234	0.142	3.190
	168	0.208	0.232	0.246	0.331	0.227	2.800
	336	0.213	0.263	0.267	0.431	0.296	2.753
	720	0.214	0.277	0.303	0.437	0.325	2.878
ETTm ₁	24	0.015	0.030	0.065	0.054	0.041	0.090
	48	0.027	0.069	0.078	0.190	0.101	0.179
	96	0.044	0.194	0.199	0.183	0.142	0.272
	288	0.103	0.401	0.411	0.186	0.318	0.462
	672	0.156	0.512	0.598	0.197	0.397	0.639
Electric.	24	0.260	0.251	0.528	0.427	0.263	0.281
	48	0.319	0.346	0.409	0.551	0.373	0.381
	168	0.427	0.544	0.959	0.893	0.609	0.599
	336	0.565	0.713	1.079	1.035	0.855	0.823
	720	0.861	1.182	1.001	1.548	1.263	1.278
Avg.		0.209	0.310	0.370	0.399	0.338	1.099

Table 2: Univariate time series forecasting results on MSE.

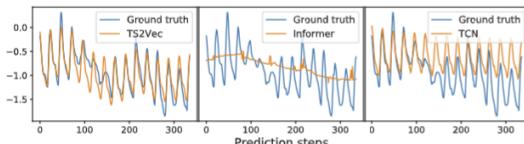


Figure 5: A prediction slice ($H=336$) of TS2Vec, Informer and TCN on the test set of ETTh₂.

Phase	H	TS2Vec	Informer			
		24	48	96	288	672
Training	24	60.42 + 2.47		402.31		
	48	60.42 + 3.63		163.41		
	96	60.42 + 5.10		392.40		
	288	60.42 + 10.76		706.94		
	672	60.42 + 21.38		938.36		
Inference	24	3.01 + 0.01		15.91		
	48	3.01 + 0.02		4.85		
	96	3.01 + 0.03		14.57		
	288	3.01 + 0.10		21.82		
	672	3.01 + 0.21		28.49		

Table 3: The running time (in seconds) comparison on multivariate forecasting task on ETTm₁ dataset.

	Yahoo			KPI		
	F ₁	Prec.	Rec.	F ₁	Prec.	Rec.
SPOT	0.338	0.269	0.454	0.217	0.786	0.126
DSPOT	0.316	0.241	0.458	0.521	0.623	0.447
DONUT	0.026	0.013	0.825	0.347	0.371	0.326
SR	0.563	0.451	0.747	0.622	0.647	0.598
TS2Vec	0.745	0.729	0.762	0.677	0.929	0.533
<i>Cold-start:</i>						
FFT	0.291	0.202	0.517	0.538	0.478	0.615
Twitter-AD	0.245	0.166	0.462	0.330	0.411	0.276
Luminol	0.388	0.254	0.818	0.417	0.306	0.650
SR	0.529	0.404	0.765	0.666	0.637	0.697
TS2Vec [†]	0.726	0.692	0.763	0.676	0.907	0.540

Table 4: Univariate time series anomaly detection results.

	Avg. Accuracy
TS2Vec	0.829
w/o Temporal Contrast	0.819 (-1.0%)
w/o Instance Contrast	0.824 (-0.5%)
w/o Hierarchical Contrast	0.812 (-1.7%)
w/o Random Cropping	0.808 (-2.1%)
w/o Timestamp Masking	0.820 (-0.9%)
w/o Input Projection Layer	0.817 (-1.2%)
<i>Positive Pair Selection</i>	
Contextual Consistency	
→ Temporal Consistency	0.807 (-2.2%)
→ Subseries Consistency	0.780 (-4.9%)
<i>Augmentations</i>	
+ Jitter	0.814 (-1.5%)
+ Scaling	0.814 (-1.5%)
+ Permutation	0.796 (-3.3%)
<i>Backbone Architectures</i>	
Dilated CNN	
→ LSTM	0.779 (-5.0%)
→ Transformer	0.647 (-18.2%)

Table 5: Ablation results on 128 UCR datasets.

- 시계열 분류 : TS2Vec는 UCR·UEA 모두에서 기존 최선 방법 대비 평균 2–3% 정확도 향상
- 예측 및 이상 탐지 : TS2Vec 기반 표현은 선형 회귀 및 간단 이상 탐지 모델에서도 성능 우수
- GPU 환경에서 학습 효율성도 좋으며, 전체 대비 학습 구조가 빠른 수렴을 가져옴

6. Insight

- 보편적 시계열 표현 : TS2Vec는 단일 모델로 세부부터 전체까지 의미 있는 표현 획득 가능
- Contrastive 구조의 장점 : 계층적 대비가 기존보다 더 다양한 의미 정보 포착
- Downstream 활용성 : 단순 선형 모형에서도 강력한 성능 제공—표현의 일반성 증명