2026 02 12
발표 자료

광운대학교 로봇학과
FAIR Lab

김한서

이번 주 진행사항

- CAIformer
    - 논문 리뷰

CAIformer

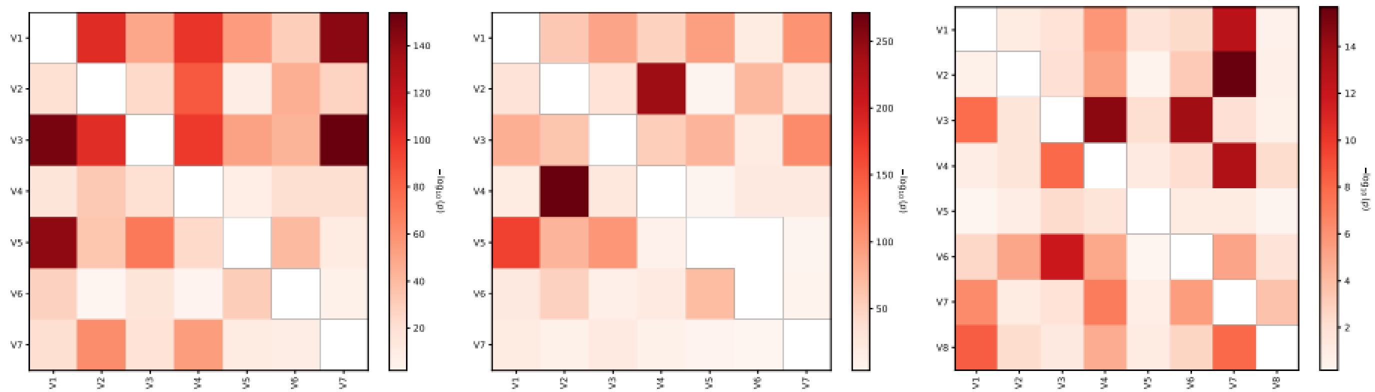
CAIFORMER: A CAUSAL INFORMED TRANSFORMER FOR MULTIVARIATE TIME SERIES FORECASTING

Xingyu Zhang[1,2], **Wenwen Qiang**[1,2], **Siyu Zhao**[1], **Huijie Guo**[1,2], **Jiangmeng Li**[1,2], **Chuxiong Sun**[1,2] **and Changwen Zheng**[1,2]
[1]University of Chinese Academy of Sciences, [2]Institute of Software, Chinese Academy of Sciences

- arXiv 등록일: 2025-05-22
- 인용 수: 3회(Google Scholar, 2026-02-10)
- ICLR 2026 reject
- 링크: https://openreview.net/pdf?id=Z97AqtNa81

CAIformer

Background



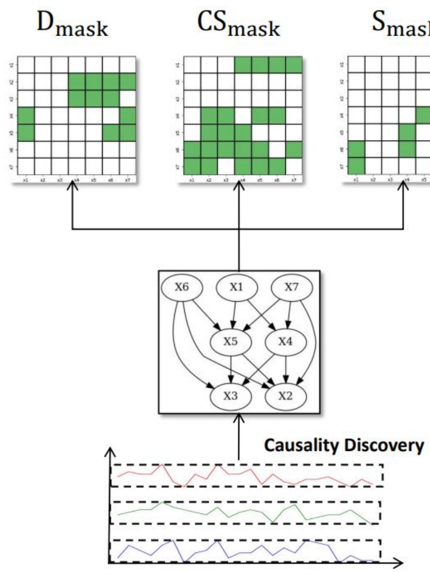
(a) ETTh1 (b) ETTm1 (c) Exchange

- 기존 Transformer 방식
  - All-to-all 예측 방식으로, 인과 관계가 없는 변수들까지 학습하기 때문에 노이즈가 발생하여 일반화 성능이 저하됨
- CAIformer 방식
  - 입력 데이터를 내생적(Endogenous), 직접 인과(Direct Causal), 콜라이더 인과(Collider Causal), 가짜 상관관계(Spurious Correlation)의 세그먼트로 분리, 이후 관계없는 변수들을 입력에서 제외하여 일반화 성능 향상

CAIformer

FAIR Lab
KWANGWOON UNIVERSITY

마스킹 생성 및 적용 과정



(c) Causal Discovery

- PC(peter-Clark) 알고리즘을 통해 변수 간 인과 관계를 나타내는 DAG(Directed Acyclic Graph) 도출
  이후에 DAG를 수치화하여 인접 행렬 $W_{adjm} \in \{-1,0,1\}^{D\times D}$ 생성
- 타겟 변수 V_i 를 기준으로 DAG 상의 경로를 분석해 4가지 Segment(ES, DCS, CCS, SCS) 정의
- 각 Segment를 인접 행렬에 적용해 D_{mask}, CS_{mask}, S_{mask} 동시 생성

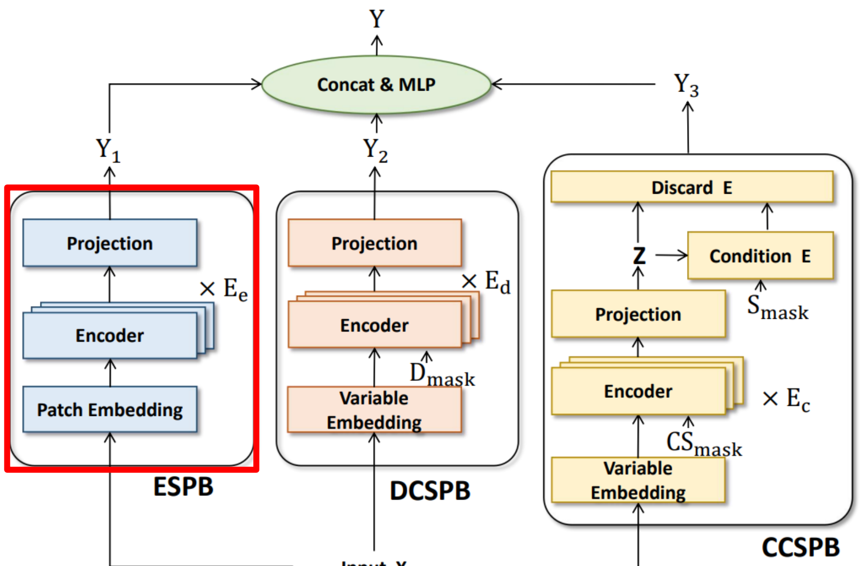
*ES (Endogenous): 타겟 변수 자신의 과거 데이터
*DCS (Direct Causal): 직접적인 원인/결과 변수
*CCS (Collider Causal): 콜라이더 구조로 얽힌 변수
*SCS (Spurious Correlation): 인과 관계가 없는 변수

CAIformer

FAIR Lab

KWANGWOON UNIVERSITY

모델 구조



(a) Framework

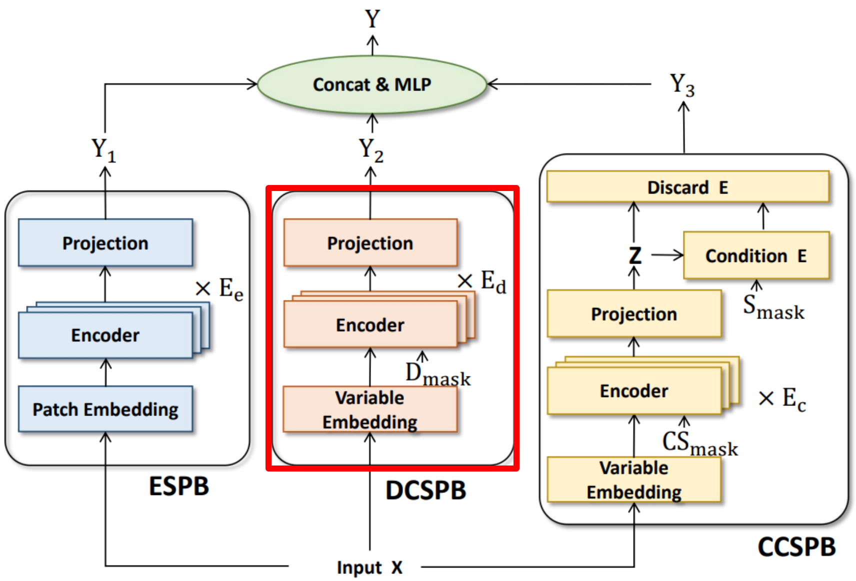
- ESPB(Endogenous Sub-segment Prediction Block)
  - Channel Independence 방식 사용
  - 외부 변수 없이, 타겟 변수 v_i 자신의 과거 데이터만을 사용하여 노이즈 방지
  - 예측 성능 변화가 변수 내부 요인 때문인지, 외부 요인 때문인지 판단하기 위해 사용

CAIformer

FAIR Lab

KWANGWOON UNIVERSITY

모델 구조



(a) Framework

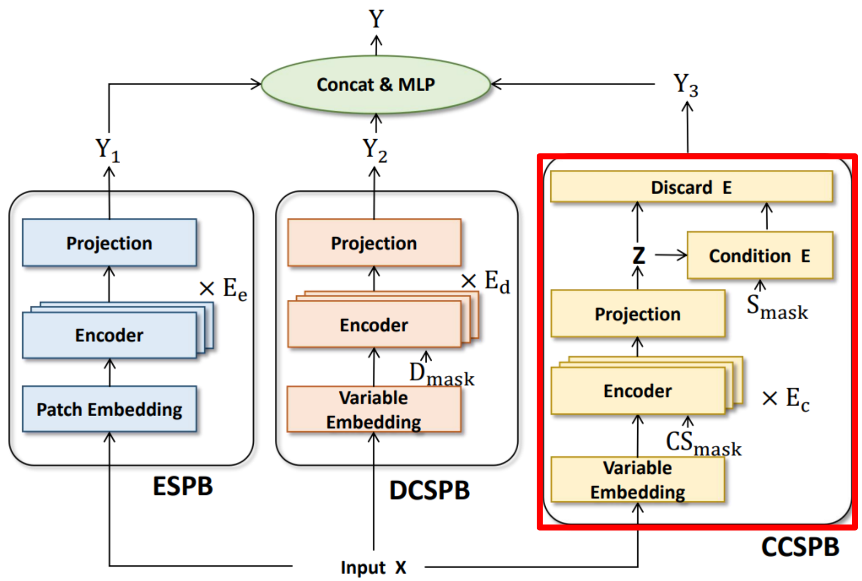
- DCSPB(Direct Causal Sub-segment Prediction Block)
    - 각 채널을 하나의 독립된 토큰으로 사용해 채널 간 상관관계 포착
    - Attention 과정에서 D_{mask}를 적용, 이를 통해 직접적인 연관이 없는 변수 간의 Attention을 방지

CAIformer

FAIR Lab

KWANGWOON UNIVERSITY

모델 구조



(a) Framework

- CCSPB(Collider Causal Sub-segment Prediction Block)
    - DCSPB와 동일하게 각 채널을 토큰화하여 사용, Attention 과정에서 CS_{mask}를 적용해 콜라이더 구조($V_i \rightarrow V_c \leftarrow V_s$)끼리만 Attention 수행
    - Condition E를 통해 콜라이더 구조에서 발생하는 가짜 상관관계를 분리 및 제거하여 일반화 성능 향상

CAIformer

FAIR Lab

KWANGWOON UNIVERSITY

주요 모델 성능 비교

Models		CAIFormer (Ours)		TimePro (2025)		SEMPO (2025)		TFPS (2025)		iTransformer (2024)		PatchTST (2023)		RLinear (2023)		Crossformer (2023)		TiDE (2023)		TimesNet (2023)		DLinear (2023)		FEDformer (2022)		Autoformer (2021)	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	96	0.327	**0.364**	**0.326**	**0.364**	0.466	0.443	0.327	0.367	0.334	0.368	0.329	0.367	0.355	0.376	0.404	0.426	0.364	0.387	0.338	0.375	0.345	0.372	0.379	0.419	0.505	0.475
	192	**0.361**	**0.377**	0.367	0.383	0.484	0.455	0.374	0.395	0.377	0.391	0.367	0.385	0.391	0.392	0.450	0.451	0.398	0.404	0.374	0.387	0.380	0.389	0.426	0.441	0.553	0.496
	336	**0.391**	**0.402**	0.402	0.409	0.506	0.469	0.401	0.408	0.426	0.420	0.399	0.410	0.424	0.415	0.532	0.515	0.428	0.425	0.410	0.411	0.413	0.413	0.445	0.459	0.621	0.537
	720	**0.449**	**0.437**	0.469	0.446	0.557	0.498	0.479	0.456	0.491	0.459	0.454	0.439	0.487	0.450	0.666	0.589	0.487	0.461	0.478	0.450	0.474	0.453	0.543	0.490	0.671	0.561
	Avg	**0.382**	**0.395**	0.391	0.400	0.503	0.466	0.395	0.406	0.407	0.410	0.387	0.400	0.414	0.407	0.513	0.496	0.419	0.419	0.400	0.406	0.403	0.407	0.448	0.452	0.588	0.517
ETTm2	96	**0.168**	**0.255**	0.178	0.260	0.196	0.286	0.170	**0.255**	0.180	0.264	0.175	0.259	0.182	0.265	0.287	0.366	0.207	0.305	0.187	0.267	0.193	0.292	0.203	0.287	0.255	0.339
	192	**0.240**	**0.302**	0.242	0.303	0.252	0.323	0.235	0.296	0.250	0.309	0.241	**0.302**	0.246	0.304	0.414	0.492	0.290	0.364	0.249	0.309	0.284	0.362	0.269	0.328	0.281	0.340
	336	0.300	0.339	0.303	0.342	0.306	0.354	**0.297**	**0.335**	0.311	0.348	0.305	0.343	0.307	0.342	0.597	0.542	0.377	0.422	0.321	0.351	0.369	0.427	0.325	0.366	0.339	0.372
	720	0.398	**0.397**	0.400	0.399	**0.391**	0.404	0.401	**0.397**	0.412	0.407	0.402	0.400	0.407	0.398	1.730	1.042	0.558	0.524	0.408	0.403	0.554	0.522	0.421	0.415	0.433	0.432
	Avg	**0.276**	0.323	0.281	0.326	0.286	0.341	0.276	**0.321**	0.288	0.332	0.281	0.326	0.286	0.327	0.757	0.610	0.358	0.404	0.291	0.333	0.350	0.401	0.305	0.349	0.327	0.371
ETTh1	96	**0.372**	0.399	0.375	0.398	0.384	0.408	0.398	0.413	0.386	0.405	0.414	0.419	0.386	**0.395**	0.423	0.448	0.479	0.464	0.384	0.402	0.386	0.400	0.376	0.419	0.449	0.459
	192	0.429	0.426	0.427	0.429	**0.409**	0.426	**0.423**	0.423	0.441	0.436	0.460	0.445	0.437	0.424	0.471	0.474	0.525	0.492	0.436	0.429	0.437	0.432	0.420	0.448	0.500	0.482
	336	0.464	0.449	0.472	0.450	**0.417**	**0.433**	0.484	0.461	0.487	0.458	0.501	0.466	0.479	0.446	0.570	0.546	0.565	0.515	0.491	0.469	0.481	0.459	0.459	0.465	0.521	0.496
	720	0.495	0.483	0.476	0.474	**0.432**	**0.454**	0.488	0.476	0.503	0.491	0.500	0.488	0.481	0.470	0.653	0.621	0.594	0.558	0.521	0.500	0.519	0.516	0.506	0.507	0.514	0.512
	Avg	0.439	0.439	0.438	0.438	**0.410**	**0.430**	0.448	0.443	0.454	0.447	0.469	0.454	0.446	0.434	0.529	0.522	0.541	0.507	0.458	0.450	0.456	0.452	0.440	0.460	0.496	0.487
ETTh2	96	0.294	0.344	0.293	0.345	**0.282**	0.342	0.313	0.355	0.297	0.349	0.302	0.348	0.288	**0.338**	0.745	0.584	0.400	0.440	0.340	0.374	0.333	0.387	0.358	0.397	0.346	0.388
	192	0.377	0.398	0.367	0.394	**0.334**	**0.384**	0.405	0.410	0.380	0.400	0.388	0.400	0.374	0.390	0.877	0.656	0.528	0.509	0.402	0.414	0.477	0.476	0.429	0.439	0.456	0.452
	336	0.425	0.430	0.419	0.431	0.355	**0.403**	0.392	0.415	0.428	0.432	0.426	0.433	**0.415**	0.426	1.043	0.731	0.643	0.571	0.452	0.452	0.594	0.541	0.496	0.487	0.482	0.486
	720	0.424	0.442	0.427	0.445	0.395	0.435	**0.410**	**0.433**	0.427	0.445	0.431	0.446	0.420	0.440	1.104	0.763	0.874	0.679	0.462	0.468	0.831	0.657	0.463	0.474	0.515	0.511
	Avg	0.380	0.403	0.377	0.403	**0.341**	**0.391**	0.380	0.403	0.383	0.407	0.387	0.407	0.374	0.398	0.942	0.684	0.611	0.550	0.414	0.427	0.559	0.515	0.437	0.449	0.450	0.459
Exchange	96	**0.083**	**0.201**	0.085	0.204	-	-	**0.083**	0.205	0.086	0.206	0.088	0.205	0.093	0.217	0.256	0.367	0.094	0.218	0.107	0.234	0.088	0.218	0.148	0.278	0.197	0.323
	192	**0.173**	**0.295**	0.178	0.299	-	-	0.174	0.297	0.177	0.299	0.176	0.299	0.184	0.307	0.470	0.509	0.184	0.307	0.226	0.344	0.176	0.315	0.271	0.315	0.300	0.369
	336	**0.292**	**0.395**	0.328	0.414	-	-	0.310	0.398	0.331	0.417	0.301	0.397	0.351	0.432	1.268	0.883	0.349	0.431	0.367	0.448	0.313	0.427	0.460	0.427	0.509	0.524
	720	0.832	0.688	**0.817**	**0.679**	-	-	1.011	0.756	0.847	0.691	0.901	0.714	0.886	0.714	1.767	1.068	0.852	0.698	0.964	0.746	0.839	0.695	1.195	0.695	1.447	0.941
	Avg	**0.345**	**0.395**	0.352	0.399	-	-	0.395	0.414	0.360	0.403	0.367	0.404	0.378	0.417	0.940	0.707	0.370	0.413	0.416	0.443	0.354	0.414	0.519	0.429	0.613	0.539
Weather	96	**0.147**	0.205	0.166	0.207	0.171	0.228	0.154	**0.202**	0.174	0.214	0.177	0.218	0.192	0.232	0.158	0.230	0.202	0.261	0.172	0.220	0.196	0.255	0.217	0.296	0.266	0.336
	192	**0.195**	**0.243**	0.216	0.254	0.218	0.269	0.205	0.249	0.221	0.254	0.225	0.259	0.240	0.271	0.206	0.277	0.242	0.298	0.219	0.261	0.237	0.296	0.276	0.336	0.307	0.367
	336	0.269	**0.285**	0.273	0.296	**0.267**	0.304	0.262	0.289	0.278	0.296	0.278	0.297	0.292	0.307	0.272	0.335	0.287	0.335	0.280	0.306	0.283	0.335	0.339	0.380	0.359	0.395
	720	0.345	**0.340**	0.351	0.346	**0.336**	0.350	0.344	0.342	0.358	0.349	0.354	0.348	0.364	0.353	0.398	0.418	0.351	0.386	0.365	0.359	0.345	0.381	0.403	0.428	0.419	0.428
	Avg	**0.239**	**0.268**	0.251	0.276	0.248	0.287	0.241	0.271	0.258	0.279	0.259	0.281	0.272	0.291	0.259	0.315	0.271	0.320	0.259	0.287	0.265	0.317	0.309	0.360	0.338	0.382
ECL	96	**0.139**	0.235	**0.139**	**0.234**	0.168	0.271	0.149	0.236	0.148	0.240	0.181	0.270	0.201	0.281	0.219	0.314	0.237	0.329	0.168	0.272	0.197	0.282	0.193	0.308	0.201	0.317
	192	**0.155**	0.245	0.156	0.249	0.183	0.283	0.162	0.253	0.162	0.253	0.188	0.274	0.201	0.283	0.231	0.322	0.236	0.330	0.184	0.289	0.196	0.285	0.201	0.315	0.222	0.334
	336	**0.171**	**0.262**	0.172	0.267	0.198	0.297	0.200	0.310	0.178	0.269	0.204	0.293	0.215	0.298	0.246	0.337	0.249	0.344	0.198	0.300	0.209	0.301	0.214	0.329	0.231	0.338
	720	**0.209**	0.303	**0.209**	**0.299**	0.238	0.329	0.220	0.320	0.225	0.317	0.246	0.324	0.257	0.331	0.280	0.363	0.284	0.373	0.220	0.320	0.245	0.333	0.246	0.355	0.254	0.361
	Avg	**0.168**	**0.261**	0.169	0.262	0.196	0.295	0.183	0.280	0.178	0.270	0.205	0.290	0.219	0.298	0.244	0.334	0.251	0.344	0.192	0.295	0.212	0.300	0.214	0.327	0.227	0.338
Traffic	96	**0.388**	**0.261**	-	-	0.441	0.333	-	-	0.395	0.268	0.462	0.295	0.649	0.389	0.522	0.290	0.805	0.493	0.593	0.321	0.650	0.396	0.587	0.366	0.613	0.388
	192	**0.411**	**0.269**	-	-	0.456	0.339	-	-	0.417	0.276	0.466	0.296	0.601	0.366	0.530	0.293	0.756	0.474	0.617	0.336	0.598	0.370	0.604	0.373	0.616	0.382
	336	**0.424**	**0.276**	-	-	0.467	0.344	-	-	0.433	0.283	0.482	0.304	0.609	0.369	0.558	0.305	0.762	0.477	0.629	0.336	0.605	0.373	0.621	0.383	0.622	0.337
	720	**0.461**	**0.297**	-	-	0.503	0.360	-	-	0.467	0.302	0.514	0.322	0.647	0.387	0.589	0.328	0.719	0.449	0.640	0.350	0.645	0.394	0.626	0.382	0.660	0.408
	Avg	**0.421**	**0.275**	-	-	0.466	0.344	-	-	0.428	0.282	0.481	0.304	0.626	0.378	0.550	0.304	0.760	0.473	0.620	0.336	0.625	0.383	0.610	0.376	0.628	0.379

Bold: 성능이 가장 좋음

Underlined: 성능이 두 번째로 좋음

모델: 최신 인과 기반 모델

CAIformer

Ablation Study

ESPB	DCSPB	CCSPB	Shuffle Mask	Weather		ETTh1		Exchange	
				MSE	MAE	MSE	MAE	MSE	MAE
w	w	w	No	0.239	0.268	0.439	0.439	0.345	0.395
w/o	w/o	w	No	0.354	0.345	0.533	0.527	0.448	0.497
w	w/o	w/o	No	0.259	0.281	0.469	0.454	0.367	0.404
w/o	w	w/o	No	0.282	0.326	0.491	0.483	0.415	0.463
w/o	w/o	w	Yes	0.378	0.369	0.551	0.548	0.471	0.522
w/o	w	w/o	Yes	0.331	0.370	0.512	0.509	0.439	0.487

- 핵심 모듈을 모두 적용했을 때의 성능 향상이 가장 큰 것을 확인하였음
  또한, ESPB 모듈의 성능 기여도가 가장 높은 것을 확인