

TimeXer: Empowering Transformers with Exogenous Variables

<https://arxiv.org/abs/2402.19072>

0. Introduction

- 시계열 예측에서 Transformer 모델이 뛰어난 성능을 보임.
- 하지만 실제 환경에서는 타겟 변수만으로는 정확한 예측이 어려움.
- 외부 요인인 외생 변수(exogenous variables)가 예측에 중요한 영향을 미침.
- 기존 모델들은 외생 변수를 동등하게 처리하거나 무시하는 경향이 있음.
- 이 논문은 외생 변수를 효과적으로 활용하는 새로운 접근법인 TimeXer를 제안함.
- TimeXer는 기존 Transformer 구조를 변경하지 않고도 외생 정보를 통합하여 예측 성능을 향상시킴.
- 실험 결과 TimeXer는 12개의 실제 시계열 예측 벤치마크에서 일관되게 우수한 성능을 보임.

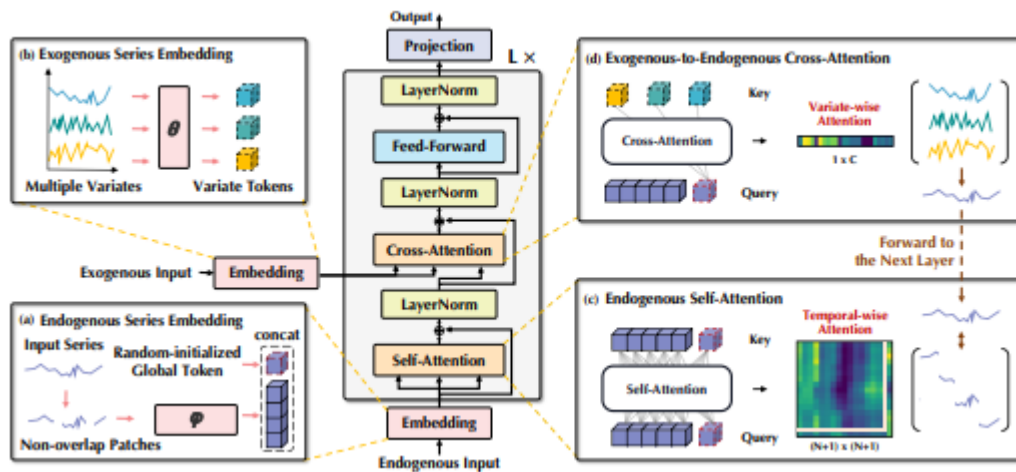
1. Overview

- TimeXer는 Transformer 기반 시계열 모델에 외생 변수 통합 기능을 추가한 방법
- 목표는 타겟 시계열만으로는 부족한 예측 정확도를 외생 변수 활용으로 개선하는 것
- 모델 구조는 기존 Transformer를 그대로 유지하고, 외생 변수 피처를 시계열 임베딩과 결합
- 외생 변수는 시계열과 시간 축을 맞춰 동기화 후, 각 시점별 입력에 포함
- TimeXer는 여러 시계열 도메인(전력, 기상, 금융 등)에서 적용 가능하도록 설계
- 구조 변경 최소화, 외생 변수 활용 극대화, 다양한 벤치마크에서 일반화 성능 확보

2. Challenges

- 시계열 예측에서 타겟 변수만으로는 계절성, 트렌드, 외부 요인 반영이 어려움
- 외생 변수 통합 시 Transformer 구조와 시간 축 정렬 문제 발생
- 외생 변수 종류가 많고 상관관계가 복잡해 학습 안정성 확보가 어려움
- 다양한 도메인과 해상도의 시계열에 일반화 가능한 모델 설계가 필요
- 기존 모델은 외생 변수 활용 시 성능 향상폭이 제한적이거나 특정 도메인에 치우침

3. Method



- TimeXer는 기존 Transformer 구조에 외생 변수 피처를 결합하는 방식으로 설계
- 시계열 임베딩과 외생 변수 임베딩을 결합해 입력 시퀀스 생성
- 외생 변수는 각 시점별로 타겟 시계열과 동기화
- Self-Attention 메커니즘으로 시계열과 외생 변수 간 상관 관계 학습
- 구조 변경 최소화, 기존 Transformer 모듈 그대로 사용 가능
- 학습 과정에서 외생 변수 중요도를 자동 조정해 과적합 방지
- 다양한 벤치마크 시계열 데이터에서 적용 가능하도록 일반화 설계

4. Experiments

- TimeXer를 12개의 실제 시계열 예측 벤치마크에서 평가

- 벤치마크 도메인 전력, 기상, 교통, 금융 등 다양한 산업 포함
- 기존 Transformer 기반 모델과 비교하여 외생 변수 통합 효과 분석
- Ablation study 진행, 외생 변수 제거, 임베딩 방식 변경, Self-Attention 수정 시 성능 확인
- 학습 효율, 일반화 성능, 장기 예측 안정성 분석
- 결과 TimeXer는 외생 변수 활용 시 모든 도메인에서 예측 정확도 향상 확인
- 특히 외생 변수와 타겟 시계열 상관관계가 높은 도메인에서 성능 개선폭 큼

5. Results

Model	TimeXer		iTransformer		RLinear		PatchTST		Crossformer		TiDE		TimesNet		DLinear		SCINet		Autoformer	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
NP	0.236	0.268	0.265	0.300	0.335	0.340	0.267	0.284	0.240	0.285	0.335	0.340	0.250	0.289	0.309	0.321	0.373	0.368	0.402	0.398
PJM	0.093	0.192	0.097	0.197	0.124	0.229	0.106	0.209	0.101	0.199	0.124	0.228	0.097	0.195	0.108	0.215	0.143	0.259	0.168	0.267
BE	0.379	0.243	0.394	0.270	0.520	0.337	0.400	0.262	0.420	0.290	0.523	0.336	0.419	0.288	0.463	0.313	0.731	0.412	0.500	0.333
FR	0.385	0.208	0.439	0.233	0.507	0.290	0.411	0.220	0.434	0.208	0.510	0.290	0.431	0.234	0.429	0.260	0.855	0.384	0.519	0.295
DE	0.440	0.415	0.479	0.443	0.574	0.498	0.461	0.432	0.574	0.430	0.568	0.496	0.502	0.446	0.520	0.463	0.565	0.497	0.674	0.544
AVG	0.307	0.265	0.335	0.289	0.412	0.339	0.330	0.282	0.354	0.284	0.412	0.338	0.340	0.290	0.366	0.314	0.533	0.384	0.453	0.368

Model	TimeXer		iTransformer		RLinear		PatchTST		Crossformer		TiDE		TimesNet		DLinear		SCINet		Autoformer	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ECL	0.171	0.270	0.178	0.270	0.219	0.298	0.205	0.290	0.244	0.334	0.251	0.244	0.192	0.295	0.212	0.300	0.268	0.365	0.227	0.338
Weather	0.241	0.271	0.258	0.278	0.272	0.291	0.259	0.281	0.259	0.315	0.271	0.320	0.259	0.287	0.265	0.317	0.292	0.363	0.338	0.382
ETTh1	0.437	0.437	0.454	0.447	0.446	0.434	0.469	0.454	0.529	0.522	0.541	0.507	0.458	0.450	0.456	0.452	0.747	0.647	0.496	0.487
ETTh2	0.367	0.396	0.383	0.407	0.374	0.398	0.387	0.407	0.942	0.684	0.611	0.550	0.414	0.427	0.559	0.515	0.954	0.723	0.450	0.459
ETTm1	0.382	0.397	0.407	0.410	0.414	0.407	0.387	0.400	0.512	0.496	0.419	0.419	0.400	0.406	0.403	0.407	0.485	0.481	0.588	0.517
ETTm2	0.274	0.322	0.288	0.332	0.286	0.327	0.281	0.326	0.757	0.610	0.358	0.404	0.291	0.333	0.350	0.401	0.571	0.537	0.327	0.371
Traffic	0.466	0.287	0.428	0.282	0.626	0.378	0.481	0.304	0.550	0.304	0.760	0.473	0.620	0.336	0.625	0.383	0.804	0.509	0.628	0.379

	Design	En.		Ex.	NP		PJM		BE		FR		DE		AVG	
					MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Cross	Ours	P+G	V		0.236	0.268	0.093	0.192	0.379	0.243	0.385	0.208	0.440	0.415	0.307	0.265
	Replace	P+G	P		0.237	0.269	0.101	0.196	0.376	0.246	0.390	0.206	0.457	0.422	0.312	0.268
	Remove	P	V		0.239	0.273	0.106	0.200	0.381	0.260	0.393	0.208	0.468	0.425	0.316	0.273
	Add	P+G	V		0.247	0.272	0.125	0.206	0.387	0.247	0.404	0.209	0.483	0.430	0.329	0.273
	Concatenate	P+G	V		0.237	0.266	0.098	0.196	0.383	0.255	0.390	0.209	0.450	0.423	0.312	0.270

Variate	Strategies	NP		PJM		BE		FR		DE		AVG	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Endogenous	Zeros	2.954	1.396	0.188	0.288	0.930	0.664	0.781	0.534	0.774	0.559	1.125	0.688
	Random	3.140	1.450	0.233	0.325	0.926	0.667	0.761	0.527	0.692	0.533	1.150	0.701
Exogenous	Zeros	0.257	0.278	0.108	0.210	0.400	0.254	0.416	0.214	0.471	0.430	0.330	0.277
	Random	0.258	0.280	0.110	0.212	0.399	0.253	0.424	0.221	0.475	0.432	0.333	0.280
TimeXer		0.236	0.268	0.093	0.192	0.379	0.243	0.385	0.208	0.440	0.415	0.307	0.265

- TimeXer는 12개 시계열 벤치마크에서 기존 Transformer 대비 평균 성능 향상
- 외생 변수 활용으로 장기 예측에서도 안정적 정확도 확보
- Ablation study에서 외생 변수 제거 시 성능 하락, 외생 변수 통합 중요성 확인
- Self-Attention과 외생 변수 결합 방식이 예측 정밀도에 큰 영향
- 도메인별 분석 전력, 금융, 교통 데이터에서 가장 큰 성능 향상, 일부 낮은 상관관계 도메인에서는 개선폭 작음
- 전체적으로 외생 변수를 효과적으로 통합한 TimeXer 구조가 시계열 예측 일반화에 기여

6. Insight

- TimeXer는 외생 변수를 효과적으로 활용해 Transformer 기반 시계열 예측 성능 향상
- 장기 예측 안정성과 일반화 능력이 개선됨
- 외생 변수 임베딩과 Self-Attention 결합 방식이 핵심
- 데이터와 도메인 특성에 따라 성능 편차 존재, 상관관계 높은 변수일수록 효과 큼
- 향후 발전 가능성으로 더 다양한 외생 변수 유형 통합, 실시간 예측, 다중 시계열 확장 연구 가능