

CaReTS : Multi-Task Time Series Forecasting Framework

<https://arxiv.org/pdf/2511.09789>

0. Introduction

- 다중 단계 시계열 예측은 에너지, 금융, 의료, 기후 등 다양한 분야에서 핵심 문제임.
- 기존 딥 러닝 기반 예측 모델은 수치 예측 정확도는 높아졌으나 추세와 편차를 분리해 해석하는 데 한계가 있음.
- 논문 기여: **CaReTS**라는 다중 태스크 프레임워크를 통해 분류(classification) + 회귀(regression)를 결합, 추세와 편차를 분리 예측함으로써 정확도와 해석 가능성 동시에 향상.

1. Overview

- **CaReTS** 프레임워크는 다중 태스크 학습 구조를 기반으로 함.
- **Dual-stream architecture:**
 - **Trend classification branch:** 미래 각 시점의 상향/하향 등의 추세 방향을 분류
 - **Regression branch:** 최신 관측값과의 편차를 회귀로 추정
- 이 구조는 거시적 의미(추세)와 미시적 변화(편차)를 분리해서 예측하므로 해석 가능성을 높임.
- 프레임워크는 CNN, LSTM, Transformer 등 다양한 Temporal encoder와 조합될 수 있는 4개의 변형(CaReTS1~4)으로 구현됨.

2. Challenges

- 시계열 예측은 미래 시간 단계가 길어질수록 정확도가 감소하는 경향이 있음.

- 대부분 모델은 단일 회귀 기반으로만 예측하여 추세 정보 해석이 어려움.
- 다중 태스크 학습 시 각 손실(task loss) 간의 균형 유지가 어려움. 이를 불확실성 기반 손실 가중치로 대응함.

3. Method

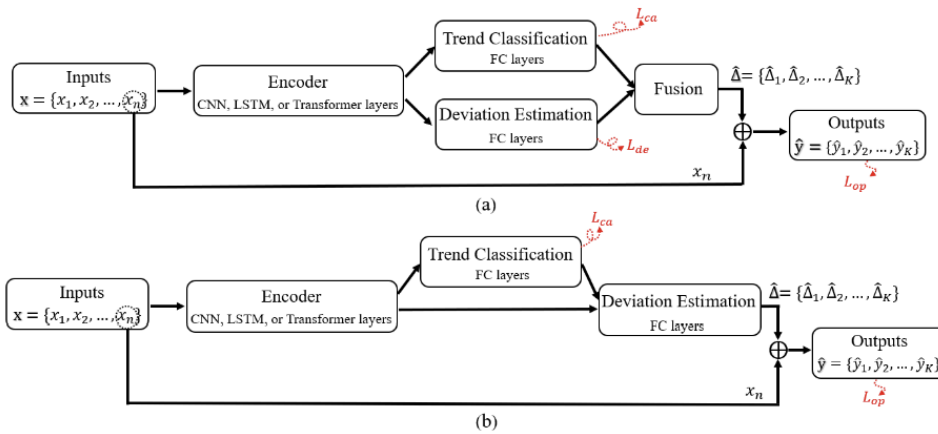


Figure 1: Two types of dual-stream CaReTS architectures

- **Dual-stream architecture**
 - Parallel streams: 추세 분류 분기와 편차 회귀 분기가 병렬로 작동
 - Sequential streams: 먼저 추세 분류 후 그 결과를 회귀 분기에 입력
- **Multi-task loss with uncertainty-aware weighting**
 - Trend classification loss, deviation regression loss, output prediction loss를 불확실성 가중치 기반으로 자동 균형 조정
- **CaReTS variants**
 - 4가지 변형(CaReTS1~4)은 각기 다른 편차/추세 처리 방식과 fusion 전략을 갖음

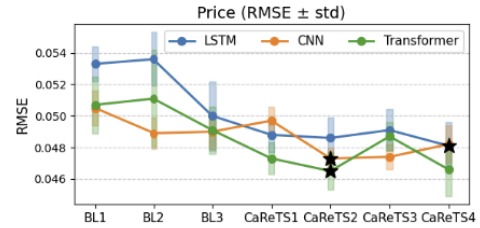
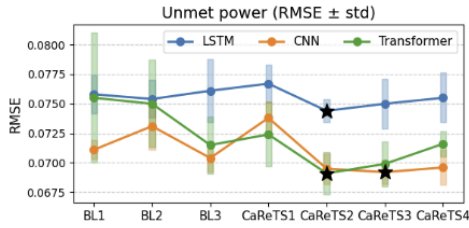
4. Experiments / Data

- **데이터셋**: 전력 수요 및 전력 가격 등 실제 연속 관측 데이터
- **설정**: 입력은 과거 관측값과 시간 정보(월, 요일, 시간 등). 예측은 미래 6단계 대상으로 수행

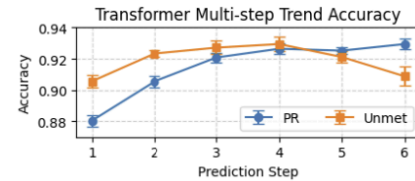
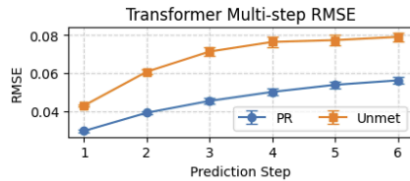
- 평가 지표: RMSE, Trend classification accuracy
- 비교 모델: SOTA(TimeXer, Autoformer, FEDformer 등) 및 전통적 예측 모델 포함

5. Results

Approach	Unmet power			Electricity price		
	Train	Validation	Test	Train	Validation	Test
LSTM						
Baseline1	0.0460 \pm 0.0040	0.0666 \pm 0.0030	0.0758 \pm 0.0016	0.0198 \pm 0.0017	0.0378 \pm 0.0027	0.0533 \pm 0.0011
Baseline2	0.0454 \pm 0.0046	0.0691 \pm 0.0016	0.0754 \pm 0.0016	0.0215 \pm 0.0019	0.0423 \pm 0.0018	0.0536 \pm 0.0017
Baseline3	0.0453 \pm 0.0049	0.0682 \pm 0.0021	0.0761 \pm 0.0027	0.0218 \pm 0.0015	0.0393 \pm 0.0013	0.0500 \pm 0.0022
CaReTS1	0.0550 \pm 0.0033	0.0723 \pm 0.0021	0.0767 \pm 0.0016	0.0265 \pm 0.0020	0.0427 \pm 0.0017	0.0488 \pm 0.0011
CaReTS2	0.0512 \pm 0.0023	0.0684 \pm 0.0025	0.0744 \pm 0.0010	0.0262 \pm 0.0019	0.0400 \pm 0.0025	0.0486 \pm 0.0013
CaReTS3	0.0474 \pm 0.0036	0.0691 \pm 0.0030	0.0750 \pm 0.0021	0.0240 \pm 0.0018	0.0408 \pm 0.0020	0.0491 \pm 0.0013
CaReTS4	0.0519 \pm 0.0032	0.0714 \pm 0.0019	0.0755 \pm 0.0021	0.0314 \pm 0.0014	0.0435 \pm 0.0024	0.0481 \pm 0.0015
CNN						
Baseline1	0.0619 \pm 0.0024	0.0679 \pm 0.0022	0.0711 \pm 0.0008	0.0410 \pm 0.0015	0.0463 \pm 0.0018	0.0505 \pm 0.0011
Baseline2	0.0550 \pm 0.0026	0.0661 \pm 0.0021	0.0731 \pm 0.0020	0.0317 \pm 0.0015	0.0408 \pm 0.0017	0.0489 \pm 0.0010
Baseline3	0.0576 \pm 0.0020	0.0655 \pm 0.0020	0.0704 \pm 0.0012	0.0310 \pm 0.0019	0.0413 \pm 0.0018	0.0490 \pm 0.0011
CaReTS1	0.0696 \pm 0.0012	0.0739 \pm 0.0029	0.0738 \pm 0.0014	0.0443 \pm 0.0013	0.0499 \pm 0.0013	0.0497 \pm 0.0009
CaReTS2	0.0658 \pm 0.0013	0.0694 \pm 0.0025	0.0695 \pm 0.0013	0.0427 \pm 0.0010	0.0470 \pm 0.0017	0.0473 \pm 0.0007
CaReTS3	0.0609 \pm 0.0020	0.0665 \pm 0.0022	0.0692 \pm 0.0010	0.0377 \pm 0.0013	0.0443 \pm 0.0015	0.0474 \pm 0.0008
CaReTS4	0.0626 \pm 0.0018	0.0678 \pm 0.0022	0.0696 \pm 0.0015	0.0428 \pm 0.0014	0.0475 \pm 0.0018	0.0482 \pm 0.0012
Transformer						
Baseline1	0.0561 \pm 0.0084	0.0683 \pm 0.0066	0.0755 \pm 0.0055	0.0322 \pm 0.0025	0.0412 \pm 0.0028	0.0507 \pm 0.0018
Baseline2	0.0530 \pm 0.0056	0.0683 \pm 0.0036	0.0750 \pm 0.0037	0.0359 \pm 0.0037	0.0436 \pm 0.0031	0.0511 \pm 0.0031
Baseline3	0.0542 \pm 0.0044	0.0667 \pm 0.0030	0.0715 \pm 0.0024	0.0353 \pm 0.0037	0.0443 \pm 0.0034	0.0491 \pm 0.0015
CaReTS1	0.0583 \pm 0.0022	0.0702 \pm 0.0028	0.0724 \pm 0.0027	0.0341 \pm 0.0016	0.0444 \pm 0.0026	0.0473 \pm 0.0010
CaReTS2	0.0588 \pm 0.0016	0.0686 \pm 0.0016	0.0691 \pm 0.0018	0.0333 \pm 0.0020	0.0445 \pm 0.0027	0.0465 \pm 0.0012
CaReTS3	0.0536 \pm 0.0026	0.0665 \pm 0.0022	0.0699 \pm 0.0019	0.0327 \pm 0.0017	0.0428 \pm 0.0028	0.0487 \pm 0.0009
CaReTS4	0.0588 \pm 0.0031	0.0696 \pm 0.0024	0.0716 \pm 0.0011	0.0375 \pm 0.0027	0.0453 \pm 0.0022	0.0466 \pm 0.0017



Encoder	CaReTS1	CaReTS2	CaReTS3	CaReTS4
Unmet power				
LSTM	0.9111 \pm 0.0020	0.9096 \pm 0.0019	0.9086 \pm 0.0030	0.9068 \pm 0.0041
CNN	0.9125 \pm 0.0032	0.9127 \pm 0.0033	0.9125 \pm 0.0032	0.9140 \pm 0.0020
Transformer	0.9191 \pm 0.0032	0.9192 \pm 0.0022	0.9168 \pm 0.0029	0.9166 \pm 0.0025
Electricity price				
LSTM	0.9073 \pm 0.0027	0.9071 \pm 0.0030	0.9066 \pm 0.0032	0.9056 \pm 0.0038
CNN	0.9032 \pm 0.0015	0.9036 \pm 0.0030	0.9024 \pm 0.0041	0.9016 \pm 0.0043
Transformer	0.9142 \pm 0.0029	0.9146 \pm 0.0019	0.9135 \pm 0.0021	0.9136 \pm 0.0051



Approach	Unmet power			Electricity price		
	RMSE	Trend Acc.	Time (s)	RMSE	Trend Acc.	Time (s)
Autoformer (Wu et al., 2021)	0.0731 \pm 0.0009	0.8891 \pm 0.0036	510.05	0.0487 \pm 0.0021	0.8713 \pm 0.0073	467.97
FEDformer (Zhou et al., 2022)	0.0908 \pm 0.0005	0.8345 \pm 0.0023	222.80	0.0874 \pm 0.0011	0.7477 \pm 0.0086	239.34
Non-stationary Transformer (Liu et al., 2022)	0.1588 \pm 0.0025	0.7384 \pm 0.0076	541.35	0.1176 \pm 0.0036	0.6970 \pm 0.0172	422.41
D-CNN-LSTM(Yao et al., 2022)	0.0732 \pm 0.0009	0.8924 \pm 0.0034	103.28	0.0573 \pm 0.0012	0.8821 \pm 0.0067	112.02
TimesNet (Wu et al., 2023)	0.0729 \pm 0.0012	0.8990 \pm 0.0028	273.15	0.0500 \pm 0.0015	0.8737 \pm 0.0074	314.40
Dlinear (Zeng et al., 2023)	0.0859 \pm 0.0004	0.8335 \pm 0.0028	68.75	0.0701 \pm 0.0005	0.7337 \pm 0.0085	70.07
Nlinear (Zeng et al., 2023)	0.1327 \pm 0.0002	0.8033 \pm 0.0017	48.44	0.1060 \pm 0.0004	0.7259 \pm 0.0036	50.33
TimeXer (Wang et al., 2024c)	0.0700 \pm 0.0022	0.9066 \pm 0.0022	448.62	0.0463 \pm 0.0013	0.9013 \pm 0.0054	573.75
TimeMixer (Wang et al., 2024a)	0.1471 \pm 0.0008	0.6983 \pm 0.0048	76.25	0.1134 \pm 0.0010	0.5831 \pm 0.0100	84.60
SOIT2FNN-MO (Yao et al., 2025b)	0.1638 \pm 0.0012	0.7021 \pm 0.0020	863.05	0.1439 \pm 0.0018	0.7153 \pm 0.0042	926.81

• 전반적 성능

- CaReTS 변형 모델들은 대부분 SOTA 대비 RMSE 성능 향상 및 추세 분류 정확도 90% 이상 달성

• Trend Classification

- CaReTS2-Transformer 등 변형이 모든 미래 단계에서 안정적으로 90% 이상의 정확도를 보임

• Ablation Study

- 다중 태스크 학습은 단일 회귀 학습 대비 RMSE 및 추세 정확도 모두 개선

• 효율성

- 계산 시간은 일부 경량 모델보다 느리지만, 대형 Transformer 기반 모델 대비 효율적

6. Insight

- **추세 & 편차 분리:** 단일 회귀 예측 대비 추세와 편차의 분리는 예측 정확도와 해석 가능성 향상에 긍정적 기여
- **불확실성 기반 손실 가중치**는 태스크 간 균형을 자동 조정해 최적화 안정성 향상
- **Framework 유연성:** 다양한 인코더(CNN/LSTM/Transformer)와 호환되어 범용성 큼
- **한계 및 확장**
 - 장기 예측(100+ steps) 성능 검증 부족
 - 테스트 데이터 도메인 다양성 제한(주로 전력 관련)
 - 보다 세분된 추세 분류 및 정량적 해석성 평가 필요