Time Series Forecasting Through the Lens of Dynamics

Alexis-Raja Brachet[1,2], Pierre-Yves Richard[2], and Céline Hudelot[1]

[1]MICS, CentraleSupélec, Université Paris-Saclay, France;
[2]CentraleSupélec, IETR UMR CNRS 6164, France

Motivations

Why do simple linear models challenge deep transformer-based ones in TSF?[1]

- From now on, analysis based on model-specific features…
- But recent SOTA foundation models are transformer-based!

We propose another perspective… Based on modality-specific features!

- For time series: **time**!

1. Zeng et al. Are Transformers Effective for Time Series Forecasting ? AAAI 2023

Learning a Dynamics!

H1: Generative models should reproduce the mechanism generating the data

- Motivated by text-based models (RNN, LSTM, Transformer, SSM,...)

H2: For time series, it is an underlying dynamics

- Motivated by PDE formalism, physics-informed technics.

H1 + H2 → **TSF models should learn a dynamics !**

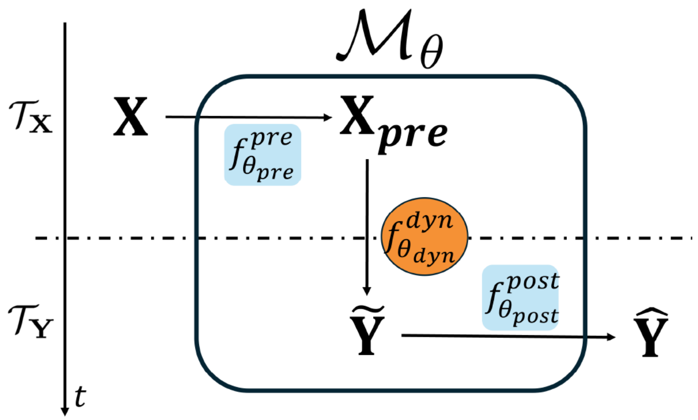
Approach

- Numerically, a TSF model maps L points to H ones.

- Idea: give a temporal meaning to computations performed along the time dimension.

The PRO-DYN nomenclature

- We propose PRO-DYN:
  - Learnable functions going forward in time = DYN functions = **model dynamics**
  - Learnable functions staying in the same time interval = PRO functions



If a non-learnable function is applied at the model entry and its inverse at the end → not considered as a PRO function (e.g., normalization, seasonal-trend decomposition,…).

An example: RNN

$$\boxed{h_{n+1} = \sigma_h(\boxed{W_{xh}}x_n + W_{hh}h_n + b_h)}$$
$$\boxed{\hat{x}_{n+1} = \sigma_y(W_{hy}h_{n+1} + b_y)}$$

The first equation:

- a **pre-PROcessing** step W_{xh}
- An **autoregressive mechanism** on h_n = DYNamics of the model: it dictates how it moves forward in time.

The second equation: **post-PROcessing** step after moving forward in time.

PRE-DYN-POST configuration

A second example: Chronos[2]

Chronos = foundation time-series model which adapts time-series modality to LLM backbones (T5, GPT-2). Processing chain:

- encoding (LLM backbone) = **pre-PROcessing** step
- next token prediction from last encoded tokens = **DYNamics step = autoregressive**

PRE-DYN configuration

2. Ansari et al. (2024). Chronos : Learning the Language of Time Series. TMLR.

Analysis

Our nomenclature aligns with:

- Success of recent transformer-based approaches[3,4]
- Fail of first transformer adaptations against LSTF-Linear models[1]

In this paper:

1. LTSF-Linear ~ discrete-time delay dynamical system
2. Better results of modified underperforming models based PRO-DYN features
3. PRE-DYN confguration (vanilla) better than DYN-POST (tested one).

1. Zeng et al. Are Transformers Effective for Time Series Forecasting ? AAAI 2023
3. Liu et al. iTransformer : Inverted transformers are effective for time series forecasting. ICLR 2024.
4. Nie et al. A time series is worth 64 words : Long-term forecasting with transformers. ICLR 2023.

Thank you!

ICML
International Conference
On Machine Learning

Interactive paper website
including 40+ analyzed models!

The paper!