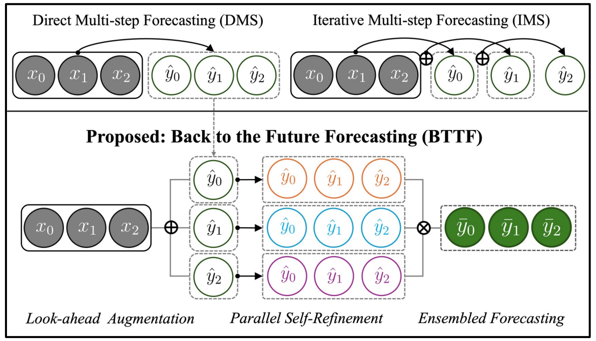
[srp518] Back to the Future: Look-ahead Augmentation and Parallel Self-Refinement for Time Series Forecasting

GitHub Paper

Sunho Kim, Susik Yoon

Korea University, Seoul, South Korea

Motivation & Key Idea



[Motivation]

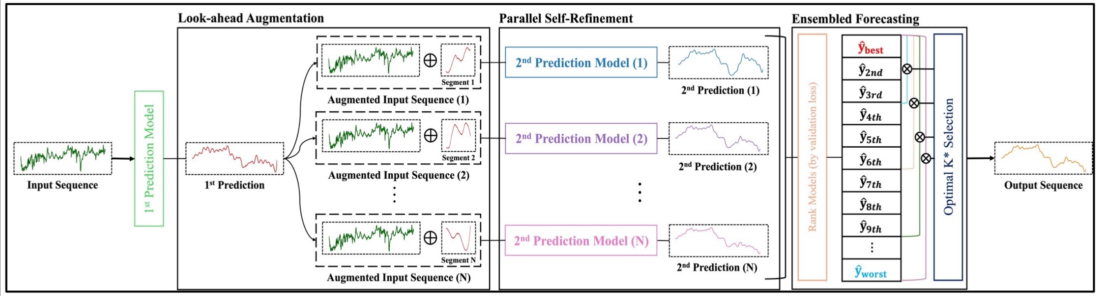
- Direct Multi-step Forecasting (DMS)
  - √ Fast parallel prediction
  - X Future temporal coherence
- Iterative Multi-step Forecasting (IMS)
  - √ Future temporal dependency
  - X Slow inference & error accumulation

Can we combine the efficiency of DMS with the temporal awareness of IMS? *Yes!*

[How?]

- Use the *model's own future predictions* as *look-ahead context* for refinement
- Refine forecasts in parallel, *without autoregressive inference*

BTTF Framework



[Main Pipeline]

1. **Initial Forecast:** Generate an initial H-step forecast

$$\hat{x}_{t+1:t+H} = f_\theta\left(x_{t-L+1}, \dots, x_t\right)$$

2. **Look-ahead Augmentation:** Split the initial forecast into future segments and append them to the input

$$\hat{x}^{(1)}_{seg,i} = \hat{x}^{(1)}_{t+s_i:t+e_i} \longrightarrow x^{(i)}_{aug} = \left[x_{t-L+1}\,;\,\hat{x}^{(1)}_{seg,i}\right], \qquad i = 1, \dots, N$$

3. **Parallel Self-Refinement:** Train segment-specific 2^{nd} stage models

$$\hat{x}^{(2)}_i = f^{(i)}_\phi\left(x^{(i)}_{aug}\right), \qquad i = 1, \dots, N$$

4. **Ensembled Forecasting:** Rank 2^{nd} stage predictions and perform incremental ensembled forecasting

$$\hat{x}^{(K)} = \frac{1}{K}\sum_{i=1}^{K} \hat{x}^{(2)}_{i-th}, \qquad K \in \{5, 10, 15, \dots, N\}$$

Results & Future Work

Linear+BTTF (ES-1E)		Linear+BTTF (1E-1E)		DLinear+BTTF (ES-1E)		DLinear+BTTF (1E-1E)		iTransformer		SegRNN	
MSE (%)	MAE (%)	MSE (%)	MAE (%)	MSE (%)	MAE (%)	MSE (%)	MAE (%)	MSE	MAE	MSE	MAE
0.7097 (56.2▲)	0.6703 (40.8▲)	0.6682 (58.7▲)	0.6547 (42.1▲)	0.6598 (6.2▲)	0.6124 (7.8▲)	0.6475 (8.0▲)	0.6189 (6.8▲)	0.6797	0.6439	0.7244	0.6280
0.6837 (20.8▲)	0.6669 (17.1▲)	0.6902 (20.1▲)	0.6725 (16.4▲)	0.7019 (11.6▲)	0.6505 (12.1▲)	0.7588 (4.5▲)	0.7070 (4.5▲)	0.7492	0.7009	0.8678	0.7119
0.7180 (15.2▲)	0.6895 (13.9▲)	0.7233 (14.6▲)	0.6970 (13.0▲)	0.7861 (8.8▲)	0.7174 (9.6▲)	0.7686 (10.8▲)	0.7266 (8.4▲)	0.8123	0.7450	0.9421	0.7596
0.7464 (15.0▲)	0.7124 (9.8▲)	0.8246 (6.1▲)	0.7601 (3.8▲)	0.7956 (17.9▲)	0.7405 (14.6▲)	0.8826 (9.0▲)	0.8046 (7.2▲)	0.9494	0.8205	0.9268	0.7768
0.1043 (6.8▲)	0.2455 (3.9▲)	0.1021 (8.7▲)	0.2458 (3.8▲)	0.1020 (3.2▲)	0.2401 (3.8▲)	0.0991 (5.9▲)	0.2415 (3.2▲)	0.1024	0.2372	0.0966	0.2279
0.2079 (16.0▲)	0.3539 (7.1▲)	0.2333 (5.7▲)	0.3827 (0.4▼)	0.2159 (29.9▲)	0.3552 (12.7▲)	0.2001 (35.0▲)	0.3535 (13.1▲)	0.2075	0.3416	0.2034	0.3353
0.3091 (27.7▲)	0.4571 (10.9▲)	0.4158 (2.8▲)	0.5071 (1.1▲)	0.3339 (-)	0.4668 (2.6▲)	0.3083 (7.7▲)	0.4544 (4.7▲)	0.4190	0.4832	0.4311	0.5001
0.8698 (25.0▲)	0.7370 (12.9▲)	0.9278 (20.0▲)	0.7596 (10.3▲)	0.8760 (29.9▲)	0.7281 (17.7▲)	0.8589 (31.3▲)	0.7232 (17.8▲)	1.1050	0.8060	1.2795	0.8724
0.0559 (4.6▲)	0.1798 (1.9▲)	0.0554 (5.5▲)	0.1797 (2.0▲)	0.0543 (6.3▲)	0.1778 (1.5▲)	0.0553 (4.6▲)	0.1781 (1.4▲)	0.0576	0.1845	0.0555	0.1838
0.0737 (4.0▲)	0.2103 (0.4▲)	0.0763 (0.7▲)	0.2094 (0.9▲)	0.0717 (1.4▲)	0.2061 (3.1▲)	0.0729 (0.3▼)	0.2066 (2.8▲)	0.0727	0.2059	0.0757	0.2182
0.0867 (11.3▲)	0.2337 (4.4▲)	0.0921 (5.7▲)	0.2383 (2.5▲)	0.0934 (6.3▲)	0.2393 (2.9▲)	0.0924 (7.3▲)	0.2391 (3.0▲)	0.0829	0.2217	0.0905	0.2399
0.1535 (4.0▲)	0.3175 (2.5▲)	0.1665 (4.1▼)	0.3322 (2.0▼)	0.1669 (10.6▲)	0.3304 (7.5▲)	0.1749 (6.3▲)	0.3444 (3.5▲)	0.0807	0.2247	0.0781	0.2257
0.0662 (-)	0.1898 (0.2▼)	0.0674 (1.7▼)	0.1917 (1.2▼)	0.0639 (0.2▼)	0.1845 (0.2▼)	0.0640 (0.3▼)	0.1843 (0.1▼)	0.0684	0.1883	0.0686	0.1891
0.0945 (1.0▼)	0.2299 (0.6▼)	0.0941 (0.6▼)	0.2299 (0.5▼)	0.0921 (0.8▼)	0.2271 (0.5▼)	0.0924 (1.2▼)	0.2275 (0.7▼)	0.1090	0.2465	0.1027	0.2377
0.1335 (8.3▼)	0.2808 (5.6▼)	0.1211 (1.8▲)	0.2642 (0.6▲)	0.1225 (3.0▼)	0.2666 (2.0▼)	0.1243 (4.5▼)	0.2673 (2.3▼)	0.1425	0.2869	0.1335	0.2783
0.1819 (2.8▼)	0.3276 (1.6▼)	0.1798 (1.6▼)	0.3274 (1.5▼)	0.1856 (6.7▼)	0.3360 (5.0▼)	0.1866 (7.2▼)	0.3321 (3.8▼)	0.1850	0.3370	0.1858	0.3347

[Experiments]

- Datasets: ILI, Exchange Rate, ETTh1, ETTm2
- Base Models: Linear, DLinear
- Baselines: iTransformer, SegRNN
- Evaluation: MSE & MAE in univariate forecasting

[Main Results]

- Improves Linear and DLinear in most settings
- Up to *58.7%* MSE reduction on ILI
- Up to *35.0%* MSE reduction on Exchange Rate
- Limited gain on ETTm2 when the base model is already well-fitted

[Future Work]

- Extend to multivariate forecasting
- Explore adaptive segment selection
- Combine heterogeneous model architectures

Contact: sunho_kim@korea.ac.kr