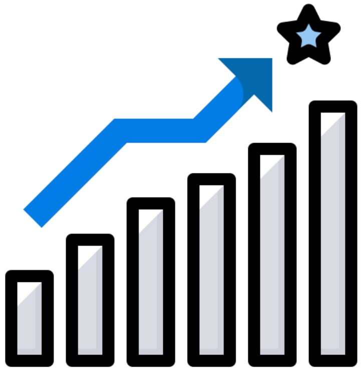
AI Project2 Final Presentation

- Stock Price Prediction Using AI Models -

2020095178 최윤선



- Table of Contents

• Introduction

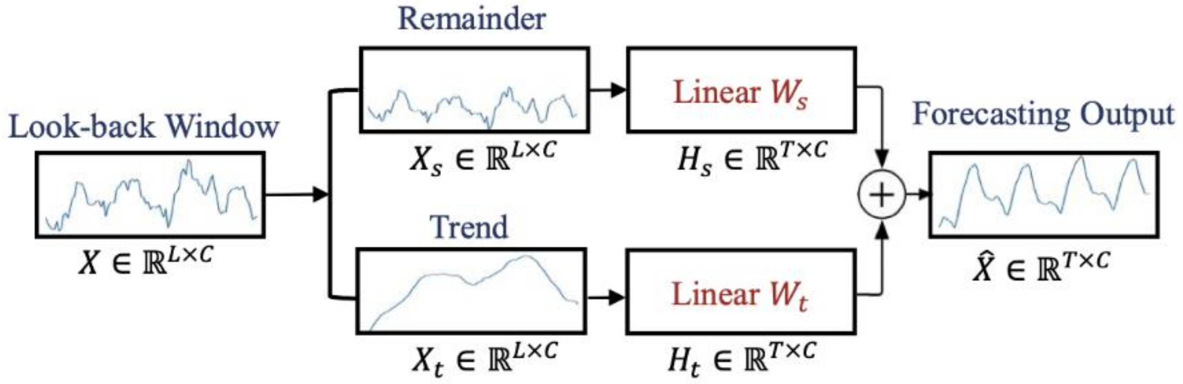
- Background
  - Train real data with some of the time series models I've theoretically studied so far
- Motivation
  - Increase in individual investors after COVID-19 → A lot of impact on the economy
  - Stock price prediction is expected to play an important role
    in establishing economic flows and financial strategies.

→ Decide to utilize **stock price data** among many time series data

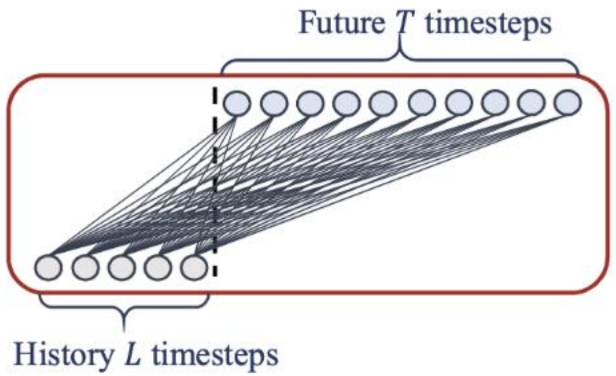
→ Comparison of predictive performance of typical time series models

DLinear, Pyraformer, PatchTST, ESTIMATE

• Model - DLinear



(a) The whole structure of *DLinear*

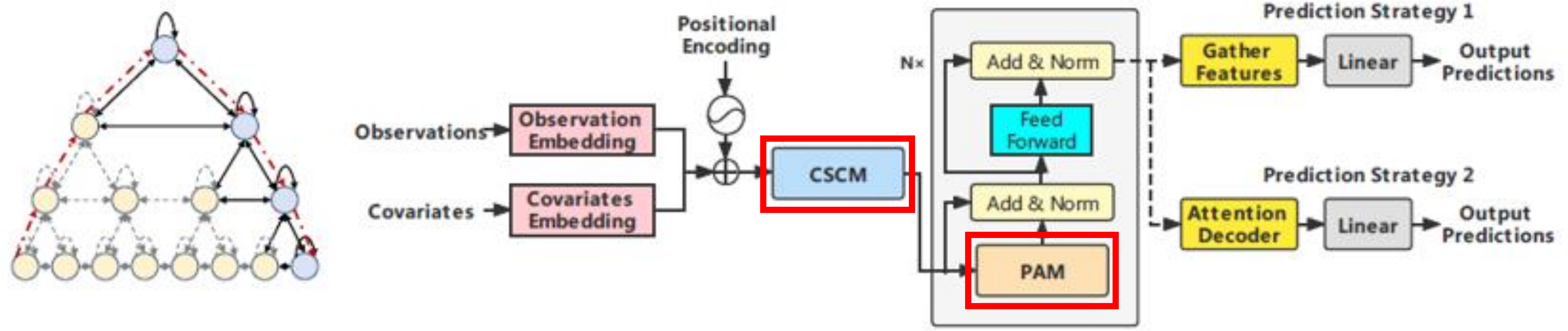


(b) One Linear Layer

- Most LTSF models have been **based on the transformer** architecture.
  → The authors raised **doubts about the effectiveness** of transformer-based models due to their **high time and memory complexity**.
- Implementation for a simpler **linear** model.
- DLinear is a **fusion of decomposition** techniques from Autoformer and FEDformer.
- Raw input data is divided into a **trend** component and a **seasonal** component.
- DLinear **boosts the performance** compared to a basic linear model.

Ailing Zeng, Muxi Chen, Lei Zhang, Qiang Xu.
Are transformers effective for time series forecasting? *arXiv preprint arXiv:2205.13504, 2022.*

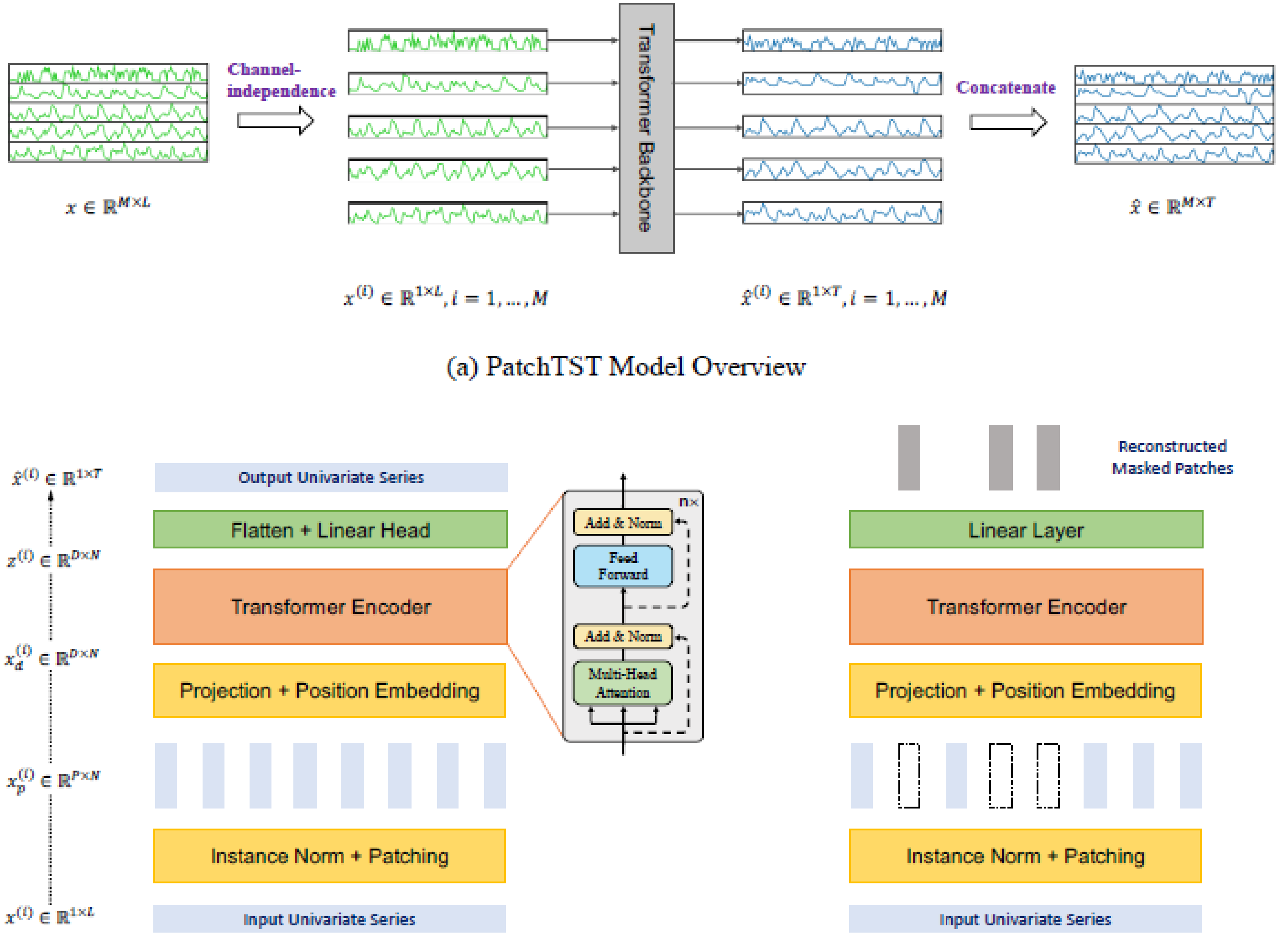
• Model - Pyraformer



- Nodes are constructed in a **pyramid shape**: Finer data to Coarser data
  - The upper nodes: information over a longer period. → Better understanding of long-range correlations.
  - The lower nodes: information over shorter periods → Better understanding of short-range correlations.
- Coarser-Scale Construction Module (**CSCM**): Form a pyramidal structure of nodes.
- Pyramidal Attention Module (**PAM**): Exchange information between the nodes of the tree passed on by CSCM.

Shizhan Liu, Hang Yu, Cong Liao, Jianguo Li, Weiyao Lin, Alex X Liu, and Schahram Dustdar.
Pyraformer: Low-complexity pyramidal attention for long-range time series modeling and forecasting. In *International Conference on Learning*

• Model - PatchTST



(a) PatchTST Model Overview

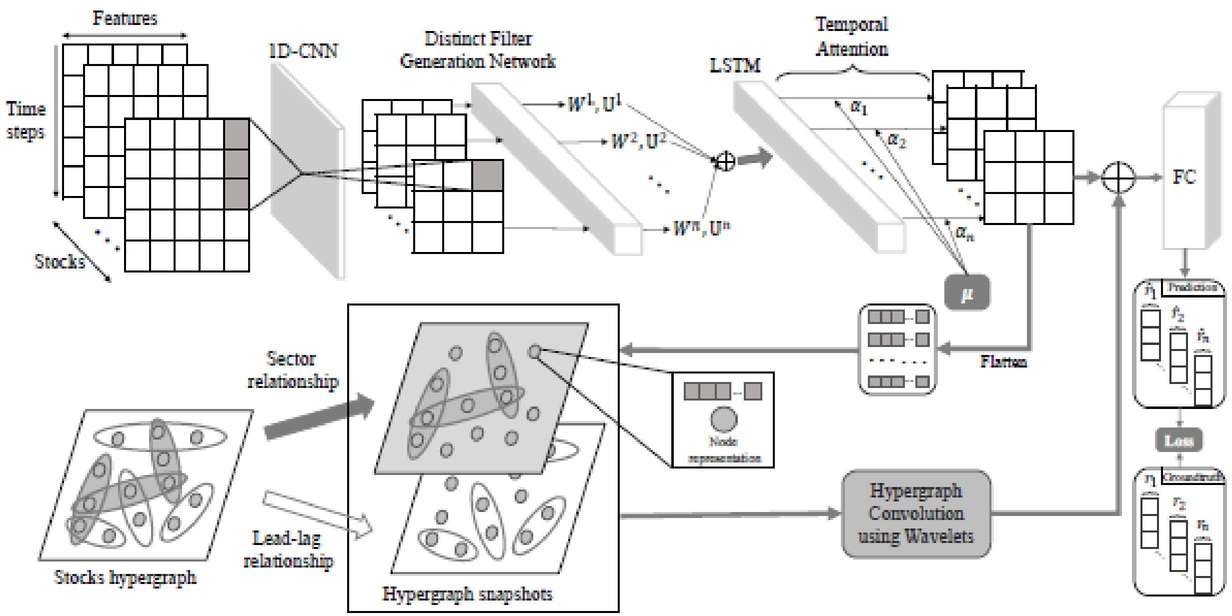
(b) Transformer Backbone (Supervised)

(c) Transformer Backbone (Self-supervised)

- Channel-independence patch
  - Each variable can be independently learned
  - Reduce the computation
  - Incorporate locality
  - Capture comprehensive semantic information
- Vanilla transformer encoder

Yuqi Nie, Nam H. Nguyen, Phanwadee Sinthong, Jayant Kalagnanam.
A time series is worth 64 words: long-term forecasting with transformers. In *International Conference on Learning Representations, 2023*

• Model - ESTIMATE



- Considerations
  - Multi-dimensional data integration(다변량), Non-stationary awareness(외부 시장 요인), Analysis of multi-order dynamics(주식 간 동적 요인), Analysis of internal dynamics(기업 내부 요인)
- Local trends – CNN → mitigate the issue of long-term dependencies
- Temporal dependencies – LSTM
- Interdependence between stock – Industry hypergraph
- Aggregate the extracted temporal information of the individual stocks – Wavelet Hypergraph Convolution

Thanh Trung Huynh, Minh Hieu Nguyen, Thanh Tam Nguyen, Phi Le Nguyen, Matthias Weidlich, Quoc Viet Hung Nguyen, Karl Aberer. Efficient integration of multi-order dynamics and internal dynamics in stock movement prediction. *arXiv preprint arXiv:2211.07400, 2022.*

- Experiment – Data

  - Data (S&P500)

	timestamp	corporation	close	open	high	low	volume	adjclose
0	2012-05-14 00:00:00	POOL	36.299999	36.360001	36.840000	35.979999	230900.0	32.237789
1	2012-05-15 00:00:00	POOL	36.750000	36.139999	37.119999	36.139999	328500.0	32.637432
2	2012-05-16 00:00:00	POOL	36.619999	36.950001	37.080002	36.549999	220300.0	32.521984
3	2012-05-17 00:00:00	POOL	34.979999	36.540001	36.540001	34.830002	418400.0	31.065506
4	2012-05-18 00:00:00	POOL	35.139999	34.900002	35.660000	34.779999	200300.0	31.207592
...	...	...	...	...	...	...	...	...
1197319	2022-05-19 00:00:00	RTX	90.250000	91.379997	92.459999	89.540001	5234700.0	90.250000
1197320	2022-05-20 00:00:00	RTX	90.080002	90.839996	91.169998	88.430000	6585500.0	90.080002
1197321	2022-05-23 00:00:00	RTX	91.830002	90.599998	92.019997	90.000000	4701300.0	91.830002
1197322	2022-05-24 00:00:00	RTX	93.209999	91.199997	93.419998	90.470001	5932000.0	93.209999
1197323	2022-05-25 00:00:00	RTX	93.559998	93.320000	94.099998	92.589996	3742000.0	93.559998

1197324 rows × 8 columns

- Training (80)
  - 2012-05-14 ~ 2020-05-20
- Validation (10)
  - 2020-05-21 ~ 2021-05-21
- Testing (10)
  - 2021-05-22 ~ 2022-05-25

• Experiment – Baseline & Evaluation

- Baseline

	Lookback Window (input sequence length)	Lookahead Window (prediction sequence length)
Short-Term Prediction	20	1
Long-Term Prediction	96	5

- Evaluation Metrics
  - MSE (Mean Squared Error), MAE (Mean Absolute Error), RMSE (Root Mean Squared Error)
  - RankIC (Rank Information Coefficient): Spearman correlation coefficient
    - Measure the rank correlation between two variables
    - (-1, 1)
  - RankIR (Rank Information Ratio): Pearson correlation coefficient
    - Quantify the linear correlation between two variables X and Y
    - (-1, 1)

• Experiment - Results

(seq, pred)	(20,1)					(96,5)				
Metrics / Models	MSE	MAE	RMSE	RankIC	RankIR	MSE	MAE	RMSE	RankIC	RankIR
DLinear	0.070	0.147	0.264	0.968	**0.974**	0.485	0.250	0.696	0.589	0.656
Pyraformer	0.072	0.123	0.268	0.946	0.876	0.087	0.138	0.295	0.922	0.846
PatchTST	**0.025**	**0.053**	**0.237**	**0.972**	0.971	**0.030**	**0.065**	**0.257**	**0.966**	**0.966**

Experiment results

Metrics / Models	MAE	RMSE	RankIC
ESTIMATE	0.053	0.572	0.034

→

Paper results

RankIC
0.516

- Comparison
  - (seq, pred) = (20, 1) → PatchTST > DLinear > Pyraformer
  - (seq, pred) = (96, 5) → PatchTST > Pyraformer > DLinear
  - (20, 1) > (96, 5)
  - Cannot compare ESTIMATE results

• Challenge

- CUDA Memory problem

  → Colab GPU, Lab Server GPU

- Because of difficulty in modifying ESTIMATE code, it is impossible to experiment under the same conditions.

  → I'm trying to adjust the length of (seq, pred) by studying the architecture of ESTIMATE

Future Work

- Modify ESTIMATE code
- Visualization of prediction results

Thank you

Q & A