CT Forecasting utilizing Time-LLM

목차

1. 시계열 예측의 발전 과정

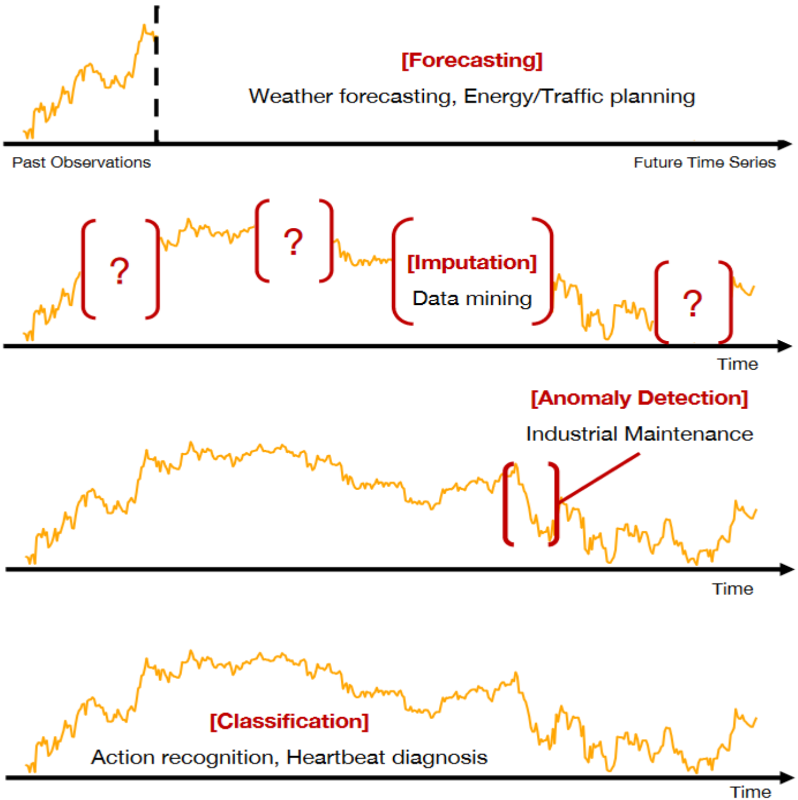
2. LLM 기반의 시계열 예측 연구 소개

3. 실제 항만 물류 분야의 적용 사례

4. 향후 연구 방향

1. 시계열 예측의 발전 과정

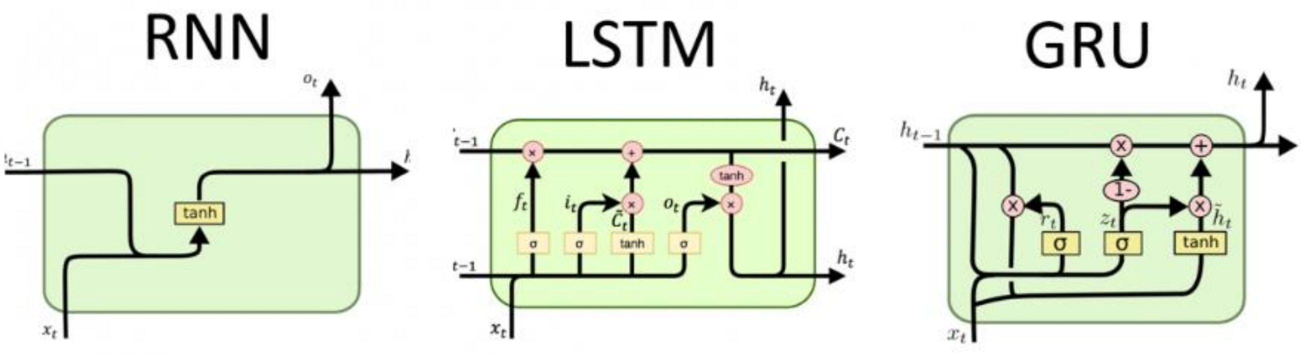
- **Time series Analysis tasks**



1. 시계열 예측의 발전 과정

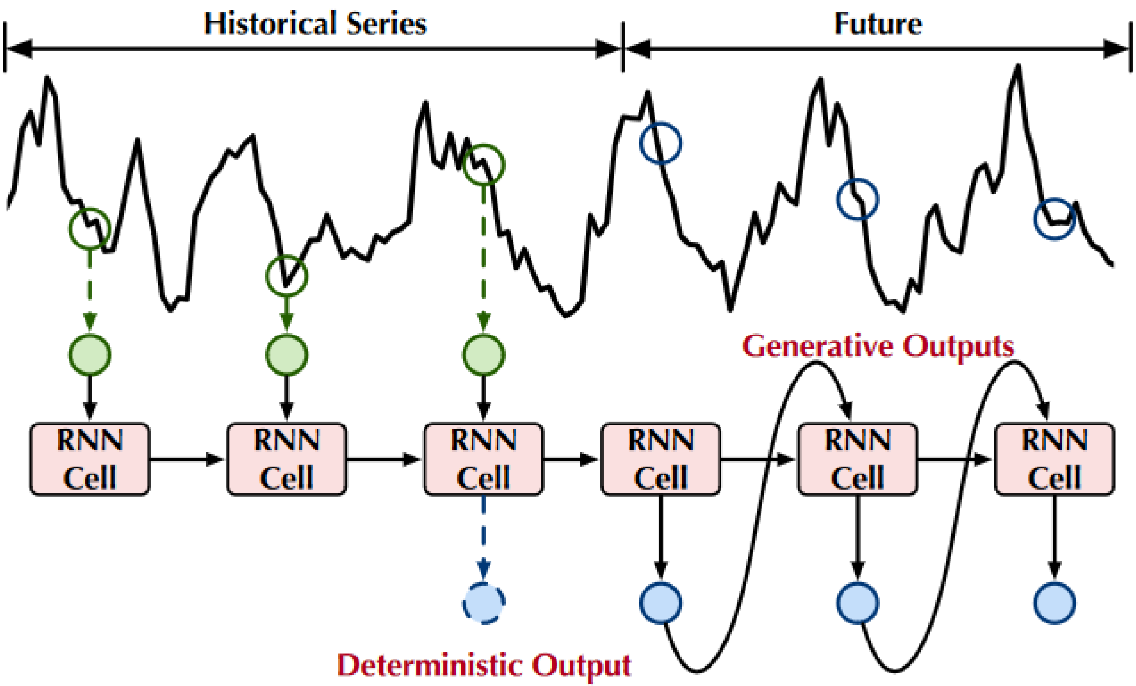
- **Traditional Approach of Time series forecasting(Recurrent)**
  - Recurrent Neural Networks (1990)
  - Long Short Term Memory (1995)
  - Bi-directional Recurrent Neural Network (1997)
  - Gated Recurrent Unit (2014)
  - Correlation Recurrent Unit (2023)

Interation recurrently of each steps with hidden state



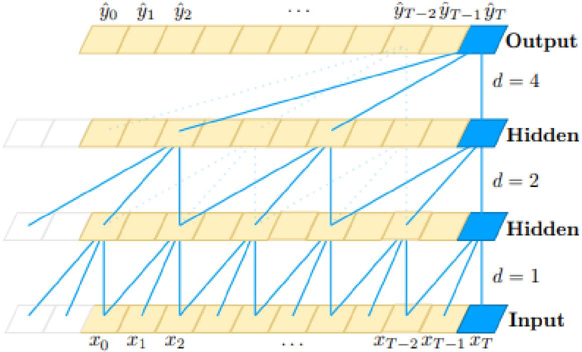
1. 시계열 예측의 발전 과정

- **Traditional Approach of Time series forecasting(Recurrent)**

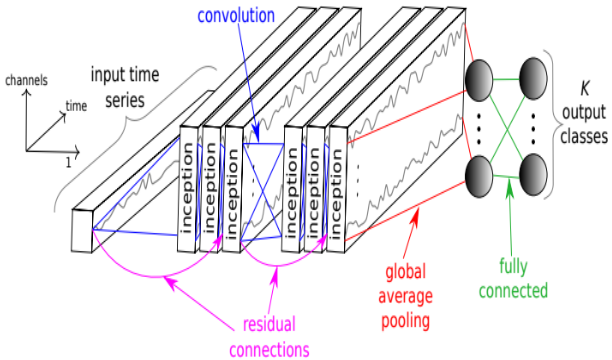


1. 시계열 예측의 발전 과정

- **Time series forecasting with Convolution Neural Network**
  - Convolutional-LSTM
  - **Temporal Convolution Network (2016)**
  - LSTNet(with RNN, 2019)
  - InceptionTime (2020)



<framework of TCN>

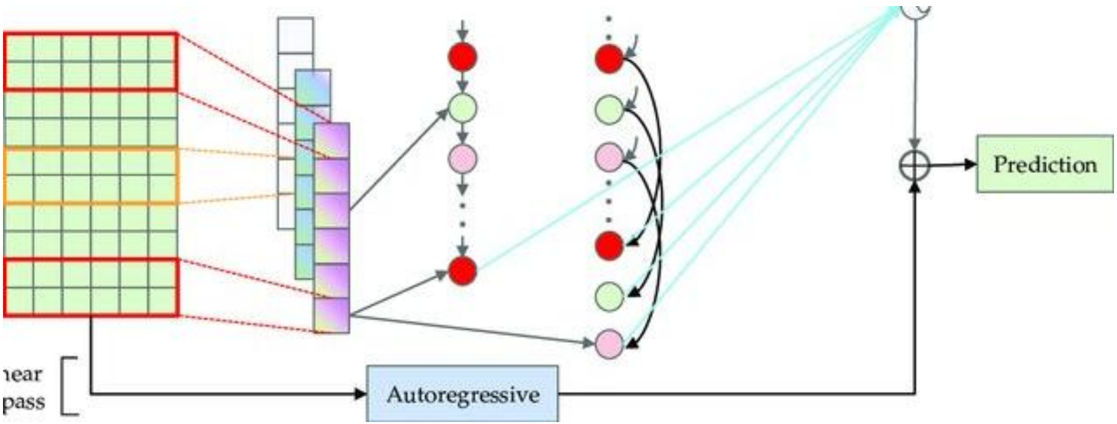


<framework of InceptionTime>

1. 시계열 예측의 발전 과정

- **Time series forecasting with Convolution Neural Network**
  - Convolutional-LSTM
  - Temporal Convolution Network (2016)
  - **LSTNet(with RNN, 2018)**
  - InceptionTime (2020)

CNN **RNN**



<Framework of LSTNet>

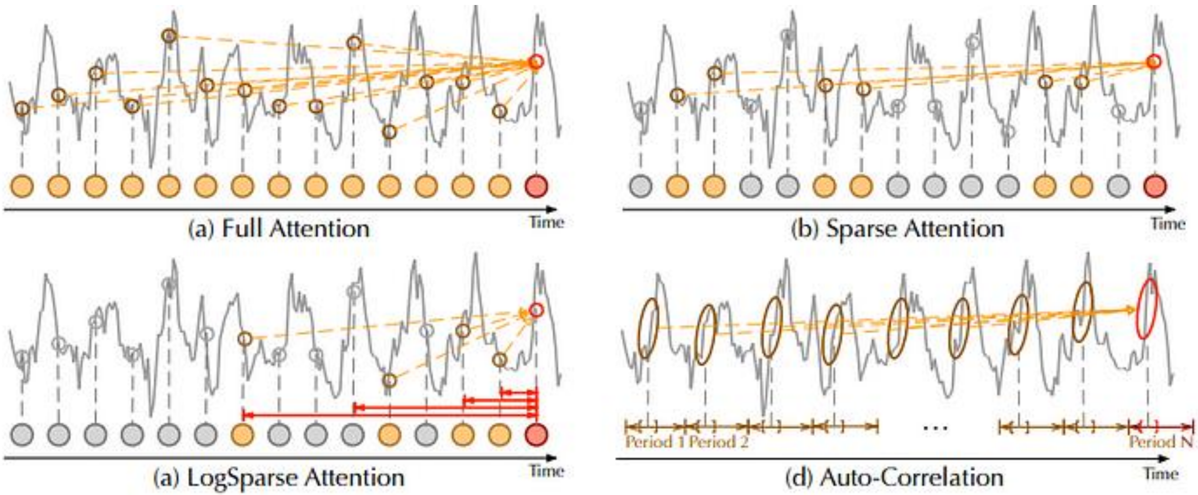
1. 시계열 예측의 발전 과정

- **Attention Approach of Time series forecasting**
  - Transformer(2017)
  - LogSparse Transformer (2019)
  - Reformer (2020)
  - Informer (2021)
  - PatchTST (2023)
  - CrossFormer (2023)

Category	Method	Architecture	Embeddding	Attention Mechanism
Vanilla	Transformer	Enc-Dec	Standard	$\mathrm{FullAttention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \mathrm{Softmax}(\frac{\mathbf{Q}\mathbf{K}^\intercal}{\sqrt{d}})$
Point-wise	LogSparse [152]	Dec-only	Standard	$\widehat{\mathbf{Q}},\widehat{\mathbf{K}} = \mathrm{CausualCov}(\mathbf{H})$ $\mathrm{FullAttention}(\widehat{\mathbf{Q}},\widehat{\mathbf{K}},\mathbf{V}) = \mathrm{Softmax}(\frac{\widehat{\mathbf{Q}}\widehat{\mathbf{K}}^\intercal}{\sqrt{d}})$
	Informer [21]	Enc-Dec	Standard	$\overline{\mathbf{Q}} = \mathrm{Top}_u\left(\overline{M}(\mathbf{q}_i,\mathbf{K})\right)$ $\mathrm{ProbSparse\text{-}Attention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \mathrm{Softmax}(\frac{\overline{\mathbf{Q}}\mathbf{K}^\intercal}{\sqrt{d}})\mathbf{V}$
	Pyraformer [153]	Enc-Dec	Standard	$\mathrm{Pyramid\text{-}Attention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \mathrm{Masked}(\mathrm{Softmax}(\frac{\mathbf{Q}\mathbf{K}^\intercal}{\sqrt{d}})\mathbf{V})$
Patch-wise	Autoformer [22]	Enc-Dec	Standard	$\mathrm{Auto\text{-}Correlation}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \sum_{i=1}^{k}\mathrm{Roll}(\mathbf{V},\tau_i)\widehat{\mathcal{R}}_{\mathbf{Q},\mathbf{K}}(\tau_i)$
	Crossformer [29]	Enc-Dec	Patch-Wise	$\mathrm{FullAttention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \mathrm{Softmax}(\frac{\mathbf{Q}\mathbf{K}^\intercal}{\sqrt{d}})$
	PatchTST [23]	Enc-only	Patch-Wise	$\mathrm{FullAttention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \mathrm{Softmax}(\frac{\mathbf{Q}\mathbf{K}^\intercal}{\sqrt{d}})$
Variate-wise	iTransformer [30]	Enc-only	Variate-Wise	$\mathrm{FullAttention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \mathrm{Softmax}(\frac{\mathbf{Q}\mathbf{K}^\intercal}{\sqrt{d}})$

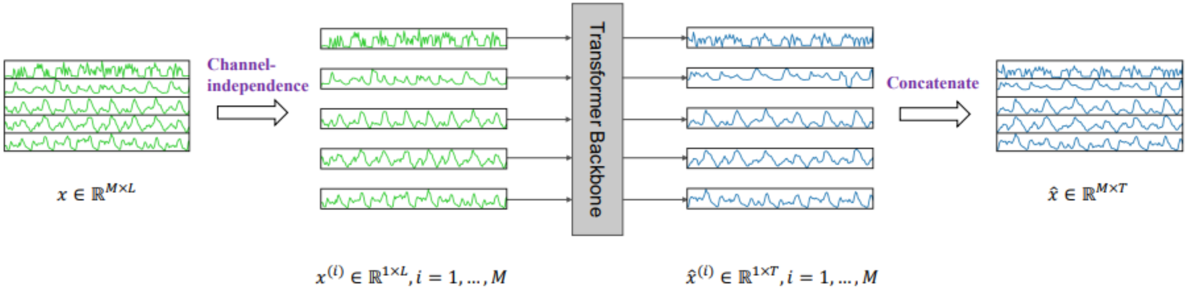
1. 시계열 예측의 발전 과정

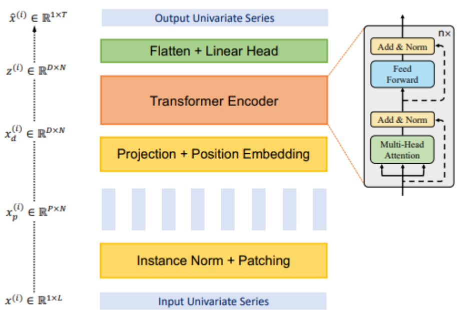
- **Attention Approach of Time series forecasting**
  - Transformer(2017)
  - LogSparse Transformer (2019)
  - Reformer (2020)
  - Informer (2021)
  - PatchTST (2023)
  - CrossFormer (2023)



1. 시계열 예측의 발전 과정

- **Attention Approach of Time series forecasting**
  - **PatchTST (2023)**

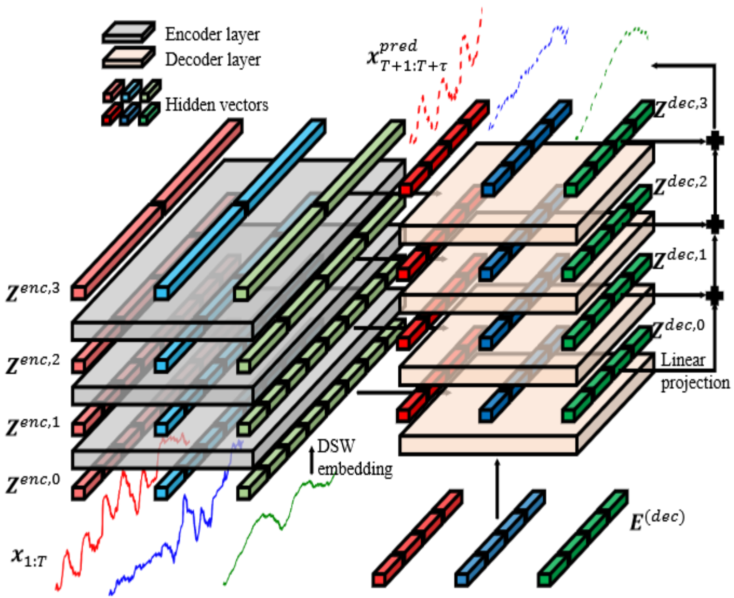


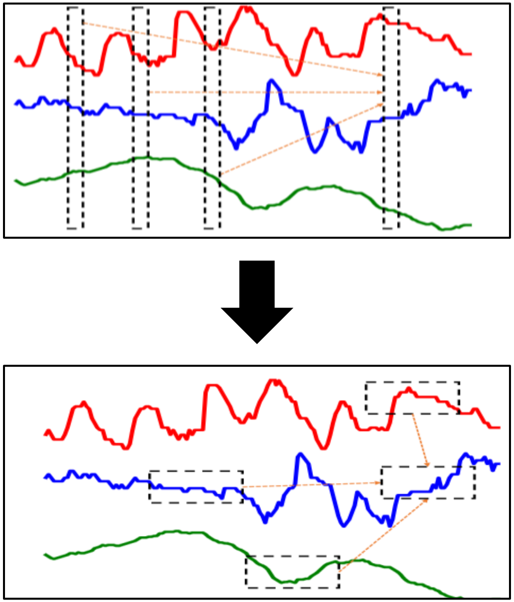


<Explicitly capturing the dependencies in multivariate time series data>

1. 시계열 예측의 발전 과정

- **Attention Approach of Time series forecasting**
  - **Crossformer (2023)**





<Explicitly capturing the dependencies in multivariate time series data>

1. 시계열 예측의 발전 과정

- **Are Transformers Effective for Time series Forecasting? (2023)**
  - **LSTF-Linear (2023)**

Are Transformers Effective for Time Series Forecasting?

Ailing Zeng[1]*, Muxi Chen[1]*, Lei Zhang[2], Qiang Xu[1]
[1]The Chinese University of Hong Kong
[2]International Digital Economy Academy (IDEA)
{alzeng, mxchen21, qxu}@cse.cuhk.edu.hk
{leizhang}@idea.edu.cn

Abstract

Recently, there has been a surge of Transformer-based solutions for the long-term time series forecasting (LTSF) task. Despite the growing performance over the past few years, we question the validity of this line of research in this

ergy management, and financial investment. Over the past several decades, TSF solutions have undergone a progression from traditional statistical methods (e.g., ARIMA [1]) and machine learning techniques (e.g., GBRT [11]) to deep learning-based solutions, e.g., Recurrent Neural Networks [15] and Temporal Convolutional Networks [3, 17].

- **Are Transformers Effective for Time series Forecasting? (2023)**
    - **LSTF-Linear (2023)**

Abstract

Recently, there has been **a surge of Transformer-based solutions** for the long-term time series forecasting (LTSF) task. Despite the growing performance over the past few years, we question the validity of this line of research in this work.

…

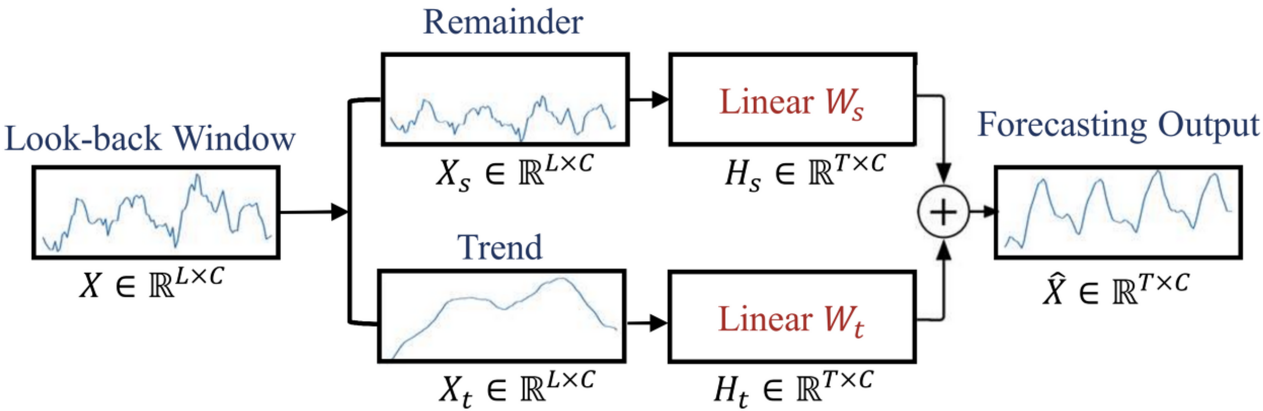
While employing positional encoding and using tokens to embed sub-series in Transformers facilitate preserving some ordering information, **the nature of the permutation-invariant self-attention mechanism inevitably results in temporal information loss**. To validate our claim, **we introduce a set of embarrassingly simple one-layer linear models** named LTSF-Linear for comparison.

…

1. 시계열 예측의 발전 과정

- **Are Transformers Effective for Time series Forecasting? (2023)**
  - ➢ **LSTF-Linear (2023)**

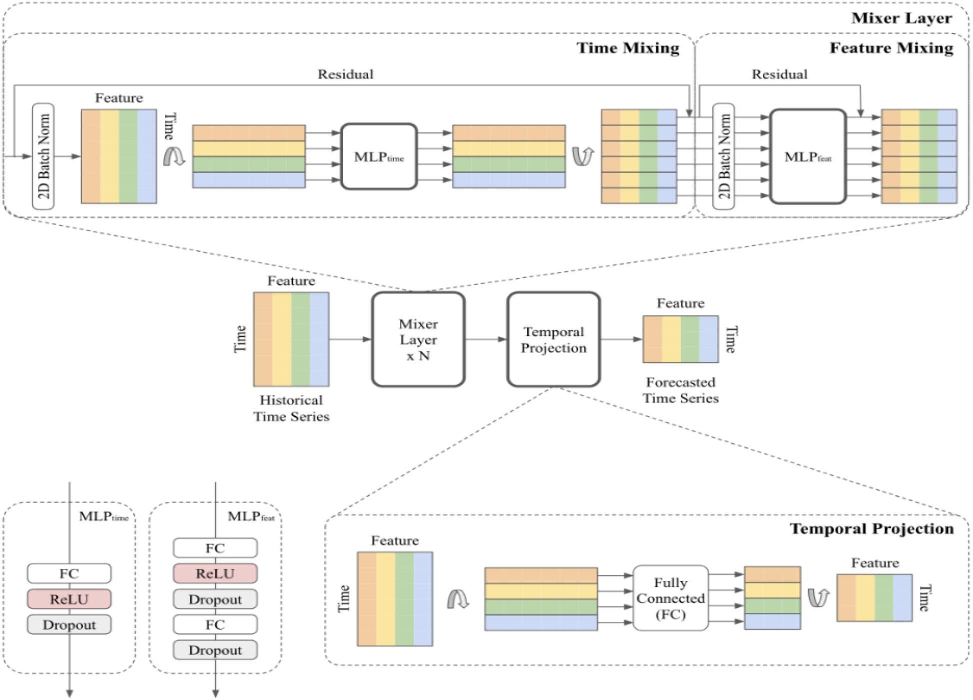
(Very Simple)



(a) The whole structure of *DLinear*

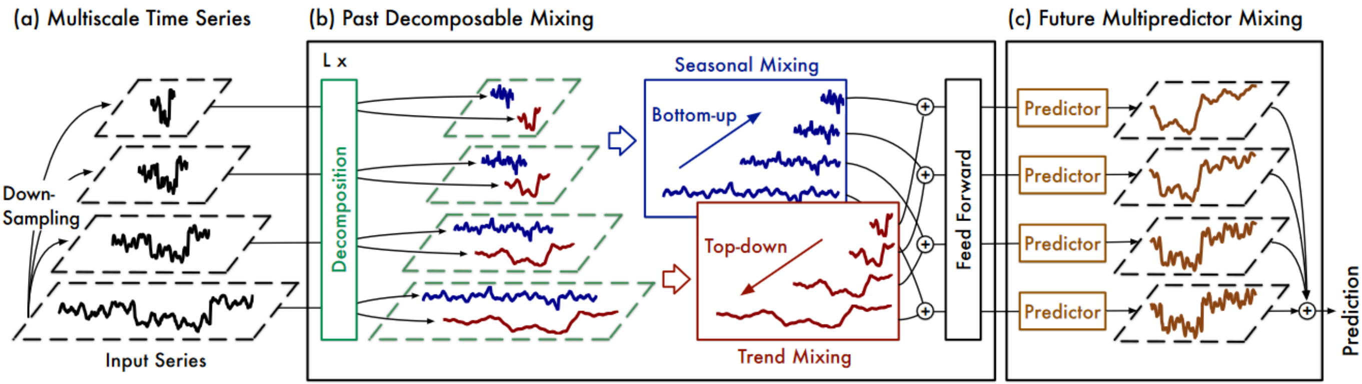
1. 시계열 예측의 발전 과정

- **MLP Approach of Time series forecasting**
  - **TSMixer (2023)**



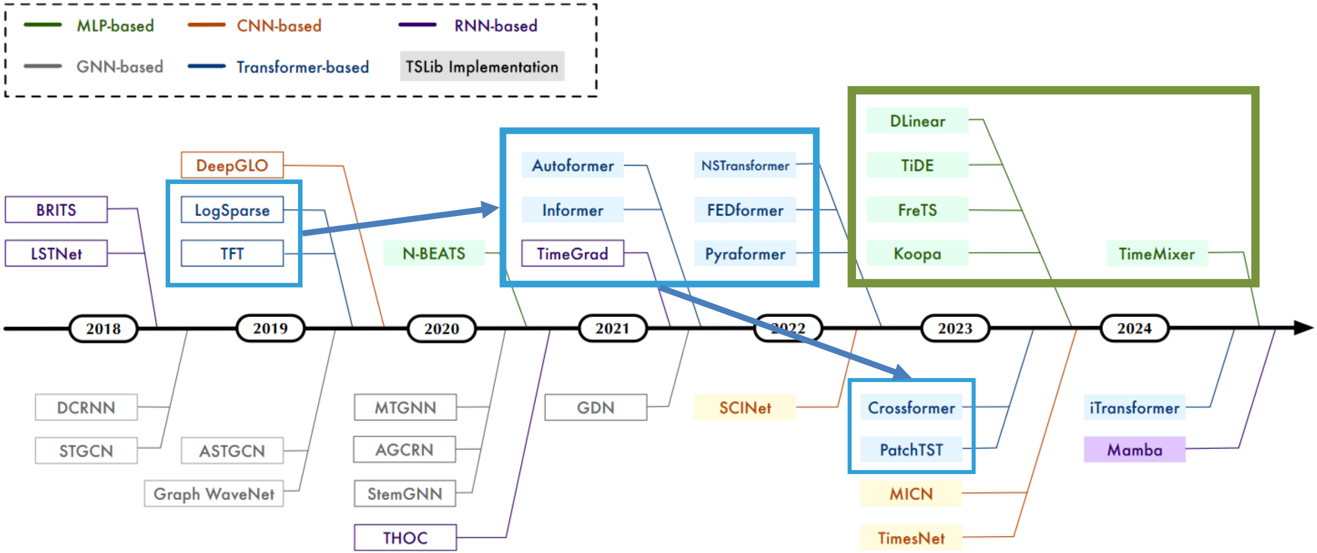
1. 시계열 예측의 발전 과정

- **MLP Approach of Time series forecasting**
  - **TimeMixer (2023)**



1. 시계열 예측의 발전 과정

- **Direction of Time series forecasting**



LLM 기반의 시계열 예측 연구 소개

2. LLM 기반의 시계열 예측 연구 소개

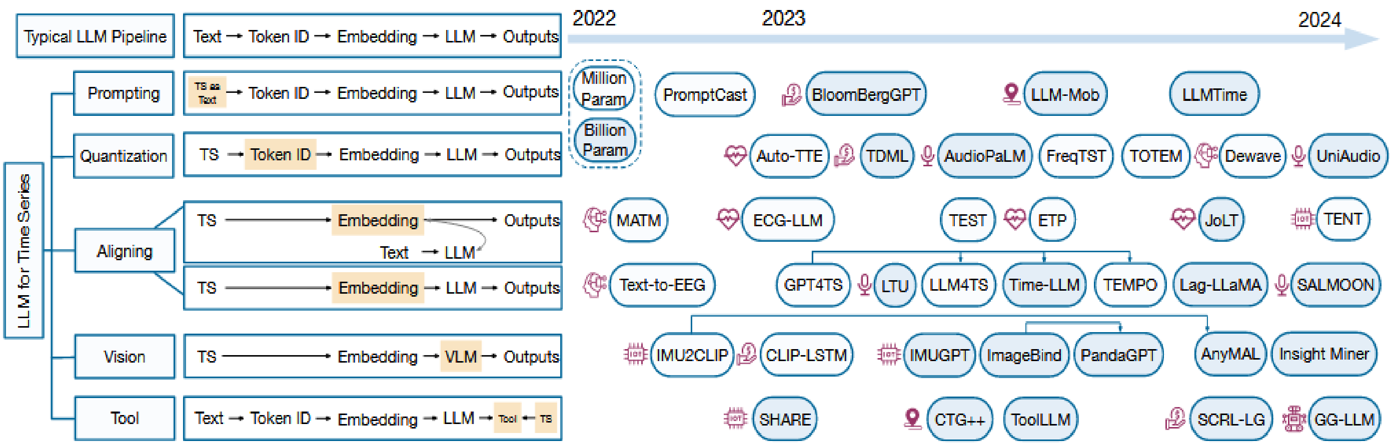
- **Why Challenging?**

Limitation	Reason	Solution
Limitation 1	LLM은 텍스트 데이터를 위해 디자인된 모델 → time series data를 representation하기에 근본적인 한계	TS Tokenization, Prompting
Limitation 2	대규모 데이터 세트의 부재 (less than 10GB) → Foundation Model 개발의 어려움	Fine-tuning
	Concept drif가 자주 발생함 → 대규모의 LLM 모델을 지속적으로 업데이트하는 것은 비효율적	
Limitation 3	text data와 time series data의 효과적인 결합(Multi Modality)의 어려움	Alignment (Multi-modal)

Zhang, X., Chowdhury, R. R., Gupta, R. K., & Shang, J. (2024). Large Language Models for Time Series: A Survey (No. arXiv:2402.01801).

2. LLM 기반의 시계열 예측 연구 소개

- **LLM 기반의 시계열 예측의 연구 분야**
  - Large Language Models for Time Series: A Survey (2024)
    - Prompting
    - Quantization(Tokenization)
    - Alignment



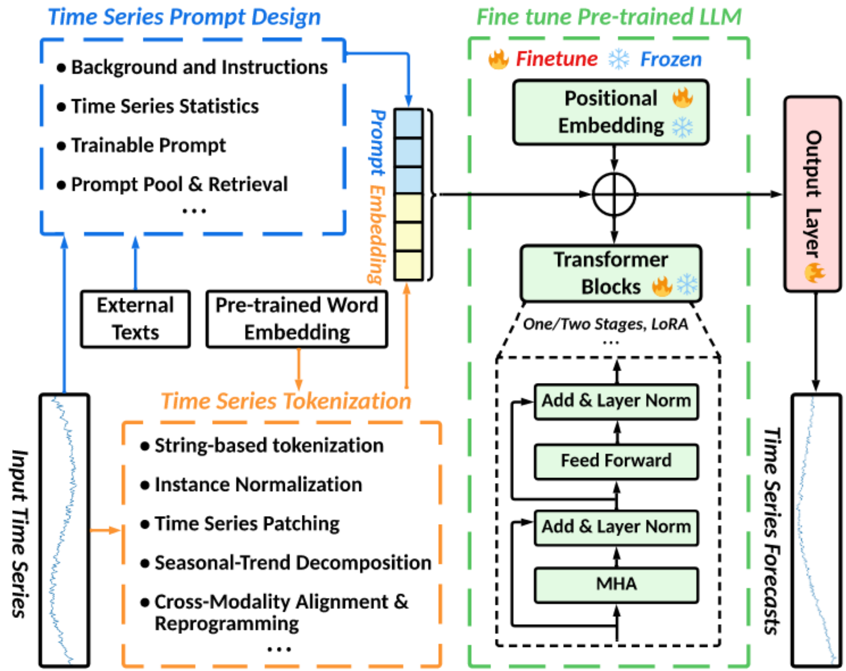
<Related Works of LLM for Time Series>

Zhang, X., Chowdhury, R. R., Gupta, R. K., & Shang, J. (2024). Large Language Models for Time Series: A Survey (No. arXiv:2402.01801).

2. LLM 기반의 시계열 예측 연구 소개

- **How we categorize various studies to develop LLM-TSF?**
  - Empowering Time Series Analysis with Large Language Models: A Survey (2024)

- **Prompting**
  - Instruction
  - Add Statistics
  - Retrieval
- **Quantization(Tokenization)**
  - String based
  - Instance Normalization
  - Patching
  - Decomposition
  - Cross-Modality
- **Fine-Tuning**



<LLM for Time Series>

Jiang, Y., Pan, Z., Zhang, X., Garg, S., Schneider, A., Nevmyvaka, Y., & Song, D. (2024). *Empowering Time Series Analysis with Large Language Models: A Survey*

2. LLM 기반의 시계열 예측 연구 소개

- **We Define 4 Categories**
- **1. Prompting**
- **2. Quantization(Tokenization)**
- **3. Alignment**
- **4. Fine-tuning**
  - ✓ 시계열 예측을 위한 **Foundation Model** 개발
  - ✓ **pre-trained LLM**을 활용하여 **fine-tuning**

2. LLM 기반의 시계열 예측 연구 소개

1. Prompting

- 다양한 연구들에서 직접 time series data를 직접 promptin하고, LLM의 지식을 바탕으로 forecasting하려는 연구 방향
- Large Language Models for Time Series: A Survey (2024)

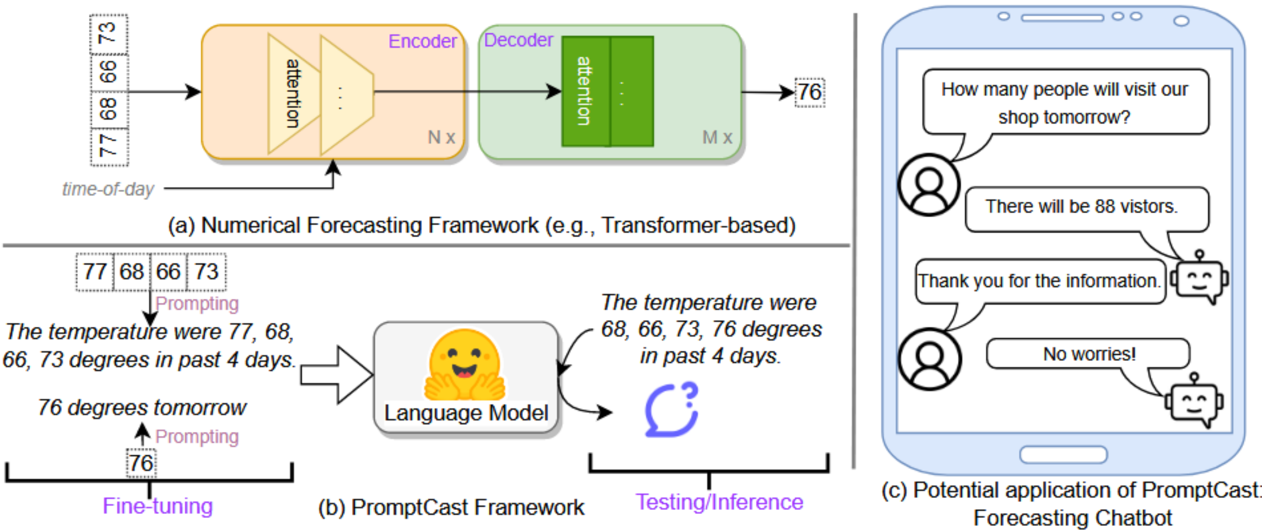
Method	Example
PromptCast [Xue and Salim, 2022]	"From $\{t_1\}$ to $\{t_{\text{obs}}\}$, the average temperature of region $\{U_m\}$ was $\{x_t^m\}$ degree on each day. What is the temperature going to be on $\{t_{\text{obs}}\}$?"
Liu *et al.* [2023d]	"Classify the following accelerometer data in meters per second squared as either walking or running: 0.052,0.052,0.052,0.051,0.052,0.055,0.051,0.056,0.06,0.064"
TabLLM [Hegselmann *et al.*, 2023]	"The person is 42 years old and has a Master's degree. She gained $594. Does this person earn more than 50000 dollars? Yes or no? Answer:"
LLMTime [Gruver *et al.*, 2023]	"0.123, 1.23, 12.3, 123.0" → "1 2 , 1 2 3 , 1 2 3 0 , 1 2 3 0 0"

Zhang, X., Chowdhury, R. R., Gupta, R. K., & Shang, J. (2024). Large Language Models for Time Series: A Survey (No. arXiv:2402.01801).

2. LLM 기반의 시계열 예측 연구 소개

1. Prompting

- **PromptCast(2023)**
  - 시계열 데이터를 문장으로 입력하고 문장으로 예측하도록 하는 sentence to sentence 기반의 Prompting 기법 제시



Xue, H., & Salim, F. D. (2023). PromptCast: A New Prompt-based Learning Paradigm for Time Series Forecasting

2. LLM 기반의 시계열 예측 연구 소개

1. Prompting

- **PromptCast(2023)**
  - 시계열 데이터를 문장으로 입력하고 문장으로 예측하도록 하는 sentence to sentence 기반의 Prompting 기법 제시

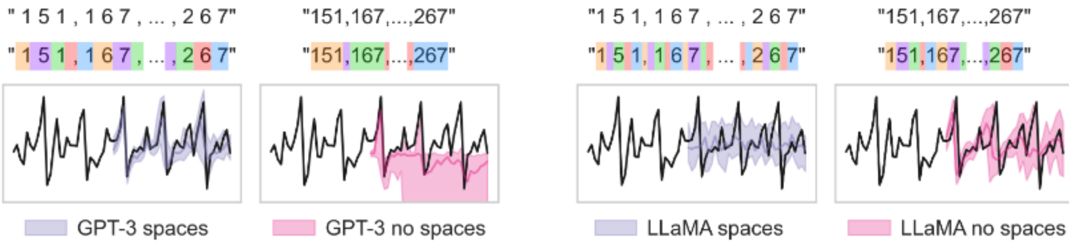
			Template	Example
CT	Input Prompt (Source)	Context	From $\{t_1\}$ to $\{t_{\text{obs}}\}$, the average temperature of region $\{U_m\}$ was $\{x^m_{t_1:t_{\text{obs}}}\}$ degree on each day.	From August 16, 2019, Friday to August 30, 2019, Friday, the average temperature of region 110 was 78, 81, 83, 84, 84, 82, 83, 78, 77, 77, 74, 77, 78, 73, 76 degree on each day.
		Question	What is the temperature going to be on $\{t_{\text{obs}+1}\}$?	What is the temperature going to be on August 31, 2019, Saturday?
	Output Prompt (Target)	Answer	The temperature will be $\{x^m_{t_{\text{obs}+1}}\}$ degree.	The temperature will be 78 degree.
ECL	Input Prompt (Source)	Context	From $\{t_1\}$ to $\{t_{\text{obs}}\}$, client $\{U_m\}$ consumed $\{x^m_{t_1:t_{\text{obs}}}\}$ kWh of electricity on each day.	From May 16, 2014, Friday to May 30, 2014, Friday, client 50 consumed 8975, 9158, 8786, 8205, 7693, 7419, 7595, 7596, 7936, 7646, 7808, 7736, 7913, 8074, 8329 kWh of electricity on each day.
		Question	What is the consumption going to be on $\{t_{\text{obs}+1}\}$?	What is the consumption going to be on May 31, 2014, Saturday?
	Output Prompt (Target)	Answer	This client will consume $\{x^m_{t_{\text{obs}+1}}\}$ kWh of electricity.	This client will consume 8337 kWh of electricity.
SG	Input Prompt (Source)	Context	From $\{t_1\}$ to $\{t_{\text{obs}}\}$, there were $\{x^m_{t_1:t_{\text{obs}}}\}$ people visiting POI $\{U_m\}$ on each day.	From May 23, 2021, Sunday to June 06, 2021, Sunday, there were 13, 17, 13, 20, 16, 16, 17, 17, 19, 20, 12, 12, 14, 12, 13 people visiting POI 324 on each day.
		Question	How many people will visit POI $\{U_m\}$ on $\{t_{\text{obs}+1}\}$?	How many people will visit POI 324 on June 07, 2021, Monday?
	Output Prompt (Target)	Answer	There will be $\{x^m_{t_{\text{obs}+1}}\}$ visitors.	There will be 15 visitors.

Xue, H., & Salim, F. D. (2023). PromptCast: A New Prompt-based Learning Paradigm for Time Series Forecasting

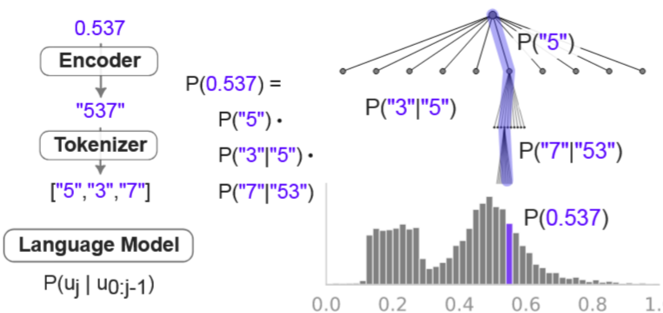
2. LLM 기반의 시계열 예측 연구 소개

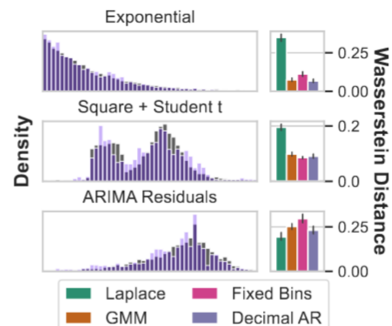
1. Prompting

- **LLMTime(2023)**
  - 숫자가 BPE Encoding 과정에서 텍스트로 변환될 때 발생하는, 정보 손실의 한계점을 극복하기 위한 Prompting 기법 제시



$$0.123, 1.23, 12.3, 123.0 \rightarrow \text{" 1 2 , 1 2 3 , 1 2 3 0 , 1 2 3 0 0"}.$$



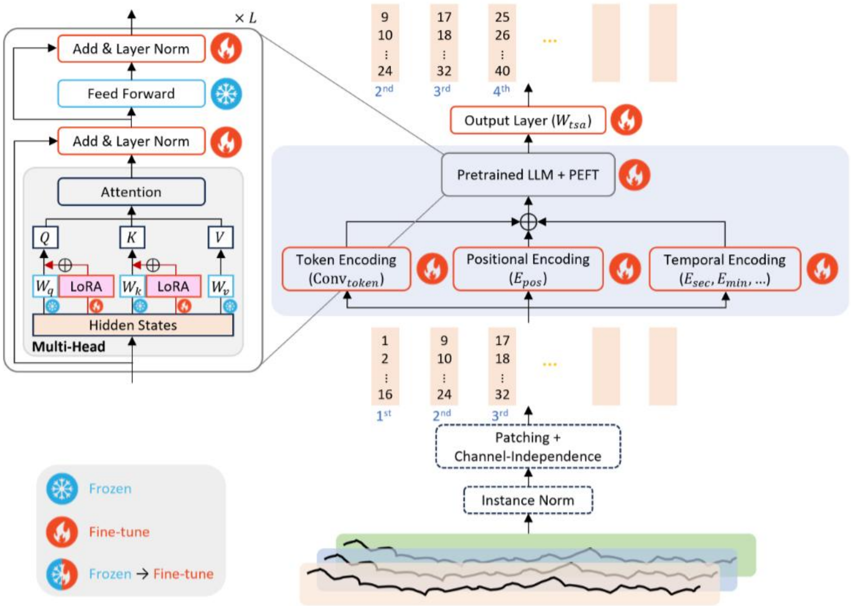


Gruver, N., Finzi, M., Qiu, S., & Wilson, A. G. (2024). LLMTime: Large Language Models Are Zero-Shot Time Series Forecasters

2. LLM 기반의 시계열 예측 연구 소개

2. Quantization(Tokenization)

- **LLM4TS(2024)**
  - 3가지 Encoding 기법(Token Encoding, Positional Encoding, Temporal Encoding) time series data를 Token형태로 변환하고 LLM에 주입



Chang, C., Wang, W.-Y., Peng, W.-C., & Chen, T.-F. (2024). LLM4TS: Aligning Pre-Trained LLMs as Data-Efficient Time-Series Forecasters

2. Quantization(Tokenization)

- **TOTEM (2025)**
  - 다양한 도메인의 time series data를 활용하여 VQ-VAE 기반 방식처럼 codebook을 생성. 이후 이 codebook 정보를 통해 다양한 downstream task를 수행

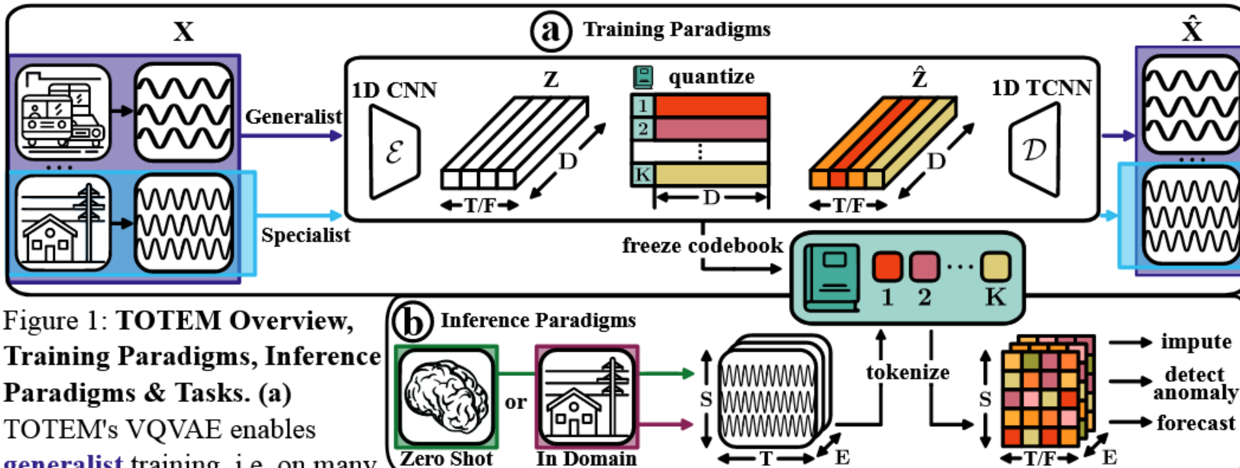


Figure 1: **TOTEM Overview, Training Paradigms, Inference Paradigms & Tasks. (a)** TOTEM's VQVAE enables **generalist** training, i.e. on many data domains jointly, and **specialist** training, i.e. on one data domain at a time. The TOTEM VQVAE architecture consists of a 1D strided CNN encoder $\mathcal{E}$, quantizer, latent codebook, and 1D strided transpose CNN decoder $\mathcal{D}$. (b) TOTEM's discrete, self-supervised codebook is frozen then leveraged for both **in domain** and **zero shot** testing across many tasks.

Talukder, S., Yue, Y., & Gkioxari, G. (2025). TOTEM: TOkenized Time Series EMbeddings for General Time Series Analysis.

3. Alignment(Multi-Modality)

- **Multimodal Pretraining of Medical Time Series and Notes**
  - 중환자실의 환자에 대한 임상 노트를 활용하여 이를 학습하고 다양한 downstream task에 적용하는 multimodal framework를 제안

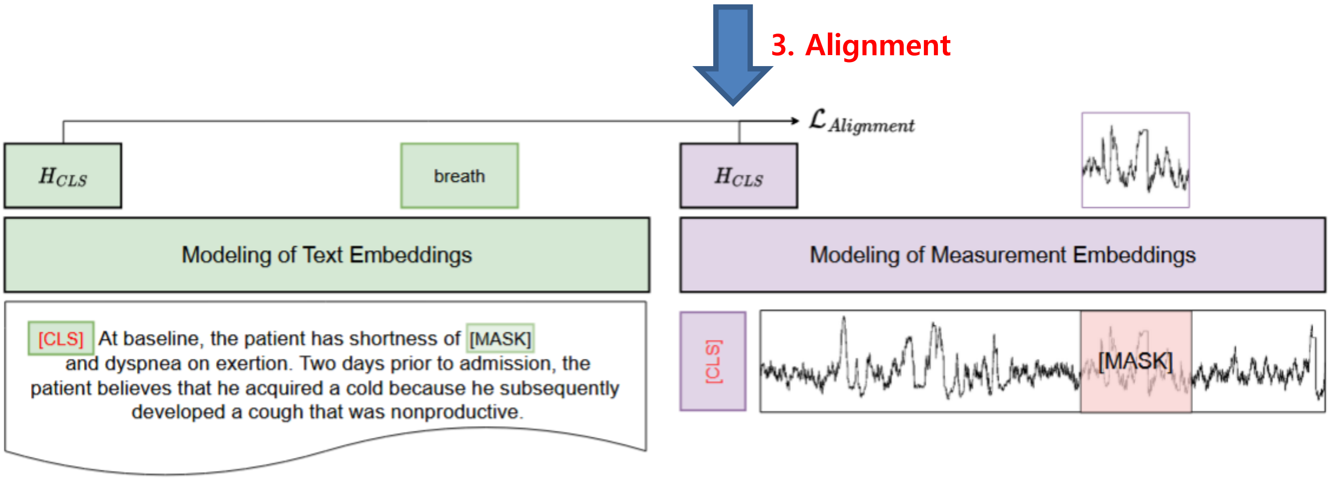


Figure 1: This figure provides a graphical representation of our proposed framework. Inputs from the different modalities are transformed by their modality specific models. The two models are trained to align the class output of the final layer, H_{CLS}, and reconstruct masked tokens.

King, R., Yang, T., & Mortazavi, B. (2023). Multimodal Pretraining of Medical Time Series and Notes

3. Alignment(Multi-Modality)

- **TimeCMA(2024)**
  - Time Series Data와 Text Prompt를 각각 Embedding하고, Multi Head Self Attention을 통해 Alingment를 수행

3. Alignment

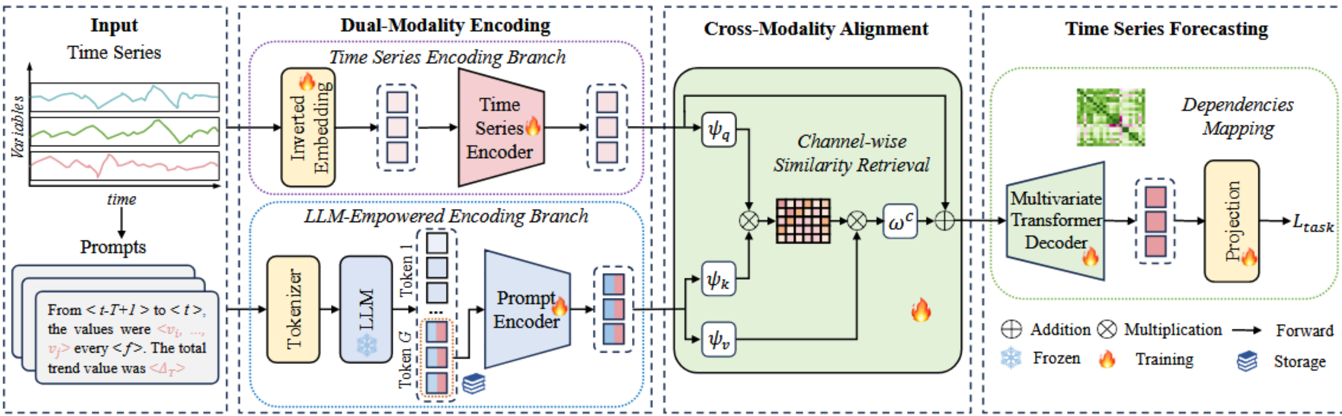
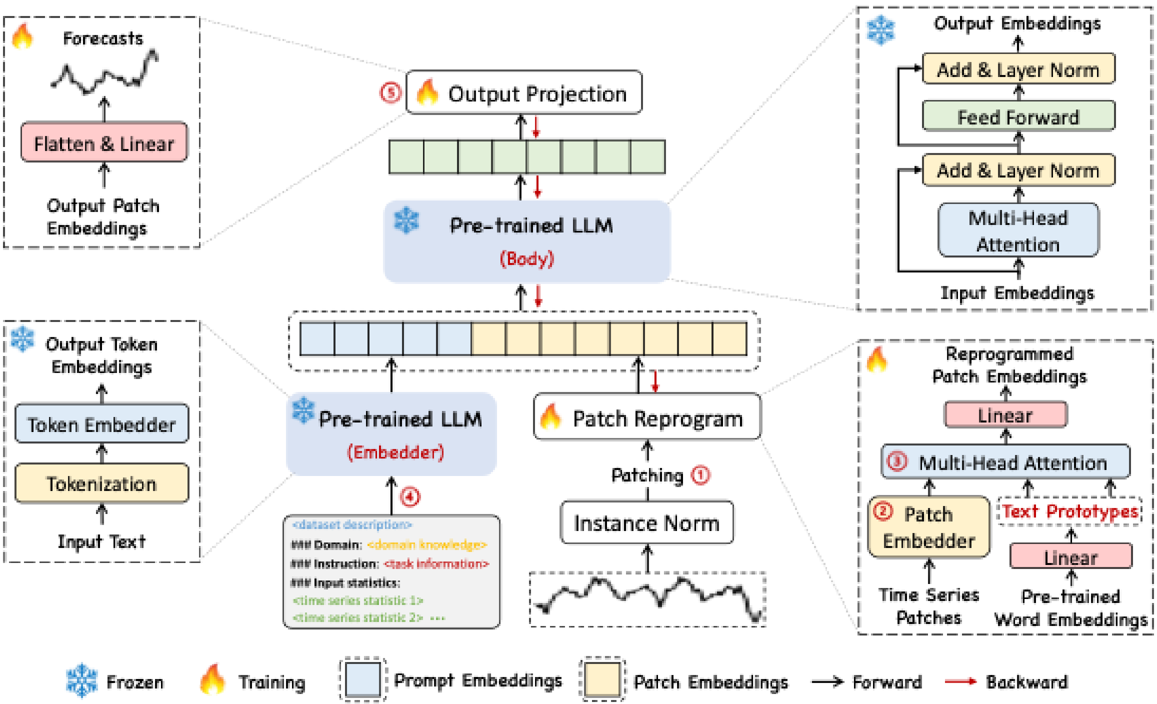


Figure 2: TimeCMA framework. Given time series data and corresponding prompts, we embed them through **Dual-Modality Encoding**. The last token G of the prompts is tailored, and its embeddings are stored for efficient inference. During **Cross-Modality Alignment**, LLM-empowered time series embeddings are retrieved from the prompt embeddings based on channel-wise similarity. **Time Series Forecasting** decodes the aligned embeddings while capturing multivariate dependencies for robust forecasting.

Liu, C., et al., (2024). TimeCMA: Towards LLM-Empowered Multivariate Time Series Forecasting via Cross-Modality Alignment

3. Alignment(Multi-Modality)

- **Time-LLM (2024)**

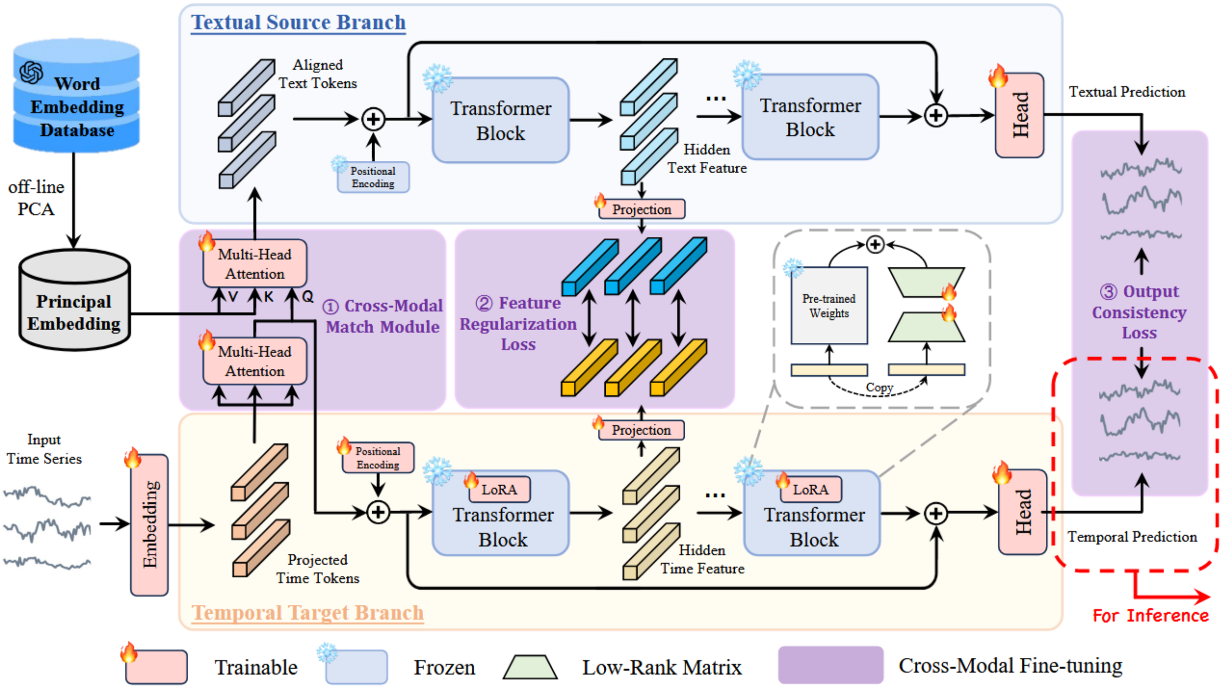


Jin, M., Wang, S., et al. (2024). Time-LLM: Time Series Forecasting by Reprogramming Large Language Models

2. LLM 기반의 시계열 예측 연구 소개

3. Alignment(Multi-Modality)

- **CALF (2024)**



Liu, P., Guo, H., Dai, T., Li, N., Bao, J., Ren, X., Jiang, Y., & Xia, S.-T. (2024). CALF: Aligning LLMs for Time Series Forecasting via Cross-modal Fine-Tuning

실제 항만 물류 분야 적용 사례

3. 실제 항만 물류 분야의 적용 사례

- **Container Throughput Forecasting**
- 항만의 컨테이너 처리량을 예측하는 항만 물류 분야의 연구 주제
- 과거의 처리량 데이터를 활용하여 미래를 예측하는 Time Series Forecasting 방법론들을 주로 채택하고 활용하고 있음

Table 1. Summary of related works

Authors	Approach	Method	Forecast Resolution Y	Q	M	D	H	Influencing factors	Forecast Step
(Yap & Lam, 2006)	MS	ECM	√					0	1
(Chou et al., 2008)	CA	Modified Regression	√					9	1
(Peng & Chu, 2009)	MS	Decomposition			√			0	1
(Schulze & Prinz, 2009)	MS	SARIMA		√				0	1
(S.-H. Chen & Chen, 2010)	MS	X-11, SARIMA			√			0	1
(J. Xiao et al., 2014)	MS	TF-DPSO			√			0	1
(Geng et al., 2015)	ML	MRSVR-CSAPSO	√					17	1
(A. Huang et al., 2015)	MS+ML	PPR-GP			√			0	1
(Patil & Sahu, 2016)	CA	Regression Model	√					2	1
(Z. Chen et al., 2016)	MS	Pearl Curve Model	√					0	1
(Pang & Gebka, 2017)	MS	SWHA, SWHM			√			0	1
(Y. Xiao et al., 2017)	MS+ML	RBFN+GRNN			√			0	1
(Xie et al., 2017)	MS+ML	DCA, LSSVR			√			0	1
(Rashed et al., 2017)	MS	ARIMA,ARIMAX			√			1	1
(M.-W. Li et al., 2017)	ML	Gauss-vSVR	√					12	1
(Tsai & Huang, 2017)	ML	ANN	√					7	1
(Niu et al., 2018)	MS+ML	VMD-ARIMA-HGWO-SVR			√			0	1
(Rashed et al., 2018)	CA	ARDL	√					6	1
(Xie et al., 2019)	MS+ML	LSSVR, ANN			√			0	1
(Shankar et al., 2019)	ML	LSTM	√	√	√			0	1
(H. Li et al., 2019)	MS+ML	SD-KELM&ELM			√			0	1~3
(Du et al., 2019)	MS+ML	VMD-BELM-ECS			√			0	1
(C.-H. Yang & Chang, 2020)	ML	CNN-LSTM			√			0	1
(Dragan et al., 2021)	MS	PCA-MLR-MC, DFM-ARIMAX		√				54	1
(Eskafi et al., 2021)	MS	Bayesian	√					6	1
(He & Wang, 2021)	ML	FGM-BP	√					0	1
(Cuong et al., 2021)	MS+ML	NPCC+AFOSMC			√			0	1
(Shankar et al., 2021)	MS+ML	ARMIAX-LSTM	√					19	1
(Milenković et al., 2021)	ML	GA-FANN			√			0	1
(Sanguri et al., 2022)	MS+ML	CTR	√	√	√			0	1
(Cuong et al., 2023)	MS+ML	DWT-LSTM			√			0	1
(Jin et al., 2023)	MS+ML	XGBoost+ARIMA				√		13	1
(Y. Xiao et al., 2023)	MS+ML	EWTTCN-KMSE			√			0	1~3
(J. Wang et al., 2024)	MS+ML	MOda, MOgoa	√					9	1
This paper	**ML**	**DPM**					√	**3**	**72**

MS: Mathematical Statistics, CA: Causal Analysis, ML: Machine Learning
Y: Yearly Q: Quarterly M: Monthly D: Daily H: Hourly

3. 실제 항만 물류 분야의 적용 사례

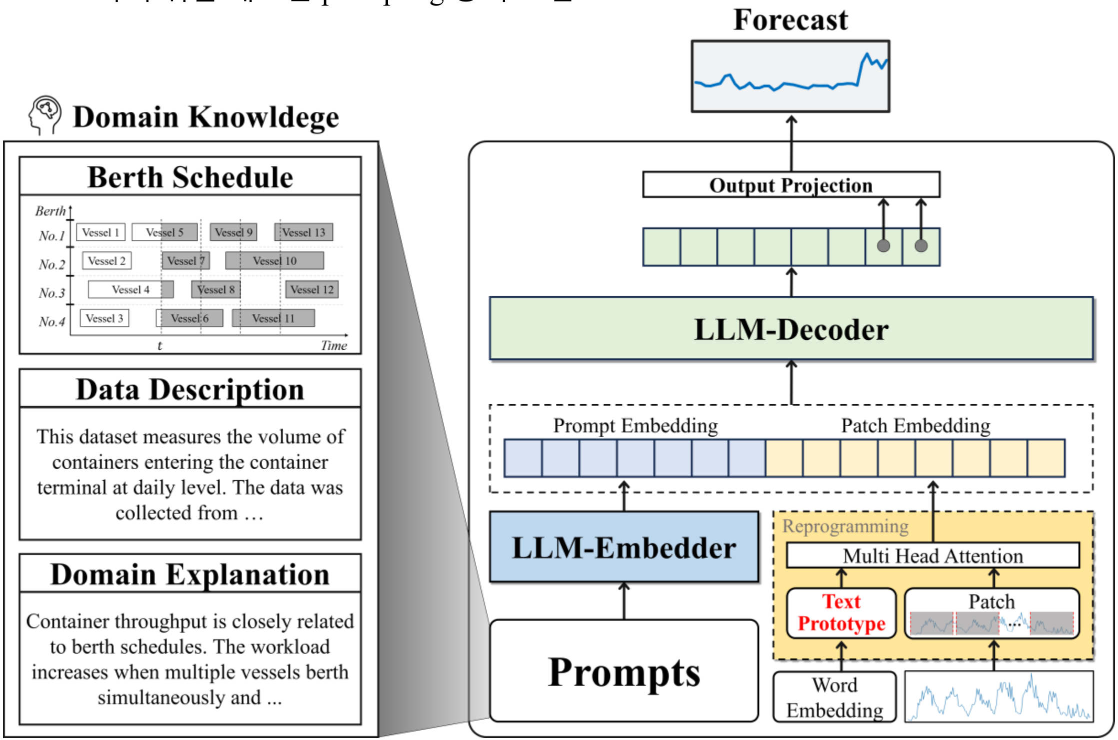
- **연구에 대한 Motivation**

- **Research Question1**: LLM을 활용한 시계열 예측 방법론이 항만 물류 분야의 task에 적용 가능 수준의 성능을 보이는가?
  - benchmark model들과의 비교를 통해 검증
  - 유의미한 성능을 보인다면, 새로운 방법론을 개발하기 위한 초석으로 활용

- **Research Question2**: 실제로 항만 물류 분야의 Domain Knowledge를 학습할 수 있는가?
  - Domain Knowledge가 학습되지 않고, 단순히 성능만 좋다면 높은 수준의 리소스를 사용하지 않는 것이 바람직함
  - 학습된 LLM에 대한 사후 분석을 통해 검증

3. 실제 항만 물류 분야의 적용 사례

- **제안 프레임워크**
  - Time-LLM 기반의 프레임워크를 채택하고 항만 물류 분야의 domain knowledg를 주입하기 위한 새로운 prompting 방식 도입



3. 실제 항만 물류 분야의 적용 사례

- 학습을 위해 활용된 Prompting 예시

[Data Description] This dataset measures the volume of containers performing gate-in/gate-out operations at a container terminal, aggregated at the daily level. The data was collected from a container terminal located at Busan Port between January 2022 and December 2022.

[Background] Container throughput is closely related to berth schedules, as increased container throughput typically corresponds to a higher volume of containers handled by arriving vessels. they have to finish their workload from ETA to ETD. And vessels scheduling to arrive in ports can identify utilizing berth schedule. Additionally, container throughput tends to decrease on weekends or public holidays because some road truck drivers do not operate during these periods.

[Task Description] [Input data] refers to the time series data, [Future data] represents the forecast period for which you need to forecast container throughput. Berth schedule will be provided to enhance predictive performance over the next <P> days.

[Input data]
Input range: 2022-01-27 ~ 2022-02-10
Weekday: [Thursday, Friday, ... Thursday]
Holiday: [Business Day, Lunar New Year, ... Business Day]
[Future data]
Forecasting range: 2022-02-10 ~ 2022-02-15
Weekday: [Thursday, Friday, ... , Tuesday]
Holiday: [Business Day, Business Day, ... , Business Day]

[Berth Schedule]
VesselCode : [A, B, C, D, ... H]
UnloadingVolume : [342, 453, 234, 464 ... 432]
LoadingVolume : [324, 123, 324, 545 ... 133]
ETA : [2022-02-09, 2022-02-10 ... 2022-02-10]
ETD : [2022-02-11, 2022-02-12 ... 2022-02-12]
Total Workload : 5432
Total Vessel : 8

3. 실제 항만 물류 분야의 적용 사례

- 다양한 Architecture를 가지고 있는 벤치마크 모델들과 비교 실험
  - Statistics, Recurrent, CNN, Transformer, Linear, LLM based model

Table 1. Experimental result

	L=14		L=21		L=28	
	P=7	P=14	P=7	P=14	P=7	P=14
Models	MSE / MAE	MSE / MAE	MSE / MAE	MSE / MAE	MSE / MAE	MSE / MAE
SARIMA	1.039 / 0.856	1.249 / 0.931	1.152 / 0.904	1.092 / 0.833	1.034 / 0.815	1.313 / 0.955
LSTM	0.884 / 0.732	1.084 / 0.865	1.021 / 0.836	1.090 / 0.887	1.189 / 0.914	1.318 / 1.005
TCN	/	/	/	/	/	/
PatchMixer	1.017 / 0.813	1.088 / 0.855	0.917 / 0.778	1.082 / 0.876	1.011 / 0.833	1.252 / 0.956
Autoformer	0.805 / 0.669	0.789 / 0.767	0.751 / 0.694	1.109 / 0.950	0.791 / 0.735	1.104 / 0.930
PatchTST	0.877 / 0.737	0.984 / 0.799	0.891 / 0.740	1.017 / 0.830	0.973 / 0.790	1.270 / 0.959
DLinear	1.580 / 1.031	1.593 / 1.043	1.750 / 1.103	1.716 / 1.103	1.555 / 1.038	2.025 / 1.217
TSMixer	1.466 / 1.000	1.715 / 1.093	1.561 / 1.051	1.587 / 1.051	2.201 / 1.206	2.405 / 1.312
Autotimes	0.893 / 0.771	0.889 / 0.776	0.831 / 0.679	0.850 / 0.702	0.634 / 0.587	0.691 / 0.629
aLLM4TS	0.735 / 0.649	0.736 / 0.626	0.606 / 0.599	0.656 / 0.627	0.599 / 0.586	0.625 / 0.633
Ours	0.653 / **0.671**	**0.614 / 0.655**	**0.566 / 0.563**	**0.569 / 0.571**	**0.530 / 0.524**	**0.565 / 0.546**
Avg	/	/	/	/	/	/
Std	/	/	/	/	/	/

3. 실제 항만 물류 분야의 적용 사례

- LLAMA를 활용하여 입력 길이/예측 구간에 따라 가장 뛰어난 성능을 보이는 backnone scale에 대한 비교 실험

Table 5. Comparison of backbone parameter

	L=14		L=21		L=28	
	P=7	P=14	P=7	P=14	P=7	P=14
Backbone	MSE / MAE	MSE / MAE	MSE / MAE	MSE / MAE	MSE / MAE	MSE / MAE
0.5 B	0.678/0.614	0.729/0.665	0.606/0.603	0.706/0.664	0.572/0.577	0.660/0.628
1.5 B	0.657/0.613	0.67/0.637	0.567/0.565	0.556/0.553	0.518/0.538	0.620/0.615
3 B	0.659/0.619	0.636/0.602	0.538/0.555	0.563/0.570	0.484/0.507	0.559/0.573
7 B	0.659/0.618	0.693/0.648	0.609/0.606	0.637/0.615	0.550/0.569	0.724/0.669
14 B	0.653/0.619	0.665/0.628	0.591/0.599	0.582/0.591	0.523/0.544	0.649/0.626

3. 실제 항만 물류 분야의 적용 사례

- 프롬프트 성능 비교 실험
  - **w/o Prompt** : 프롬프트 없이 학습
  - **w/ Prompt** : Time-LLM에서 제공하는 Prompt 정보만 활용
  - **+Domain** : 제안된 Prompt를 활용

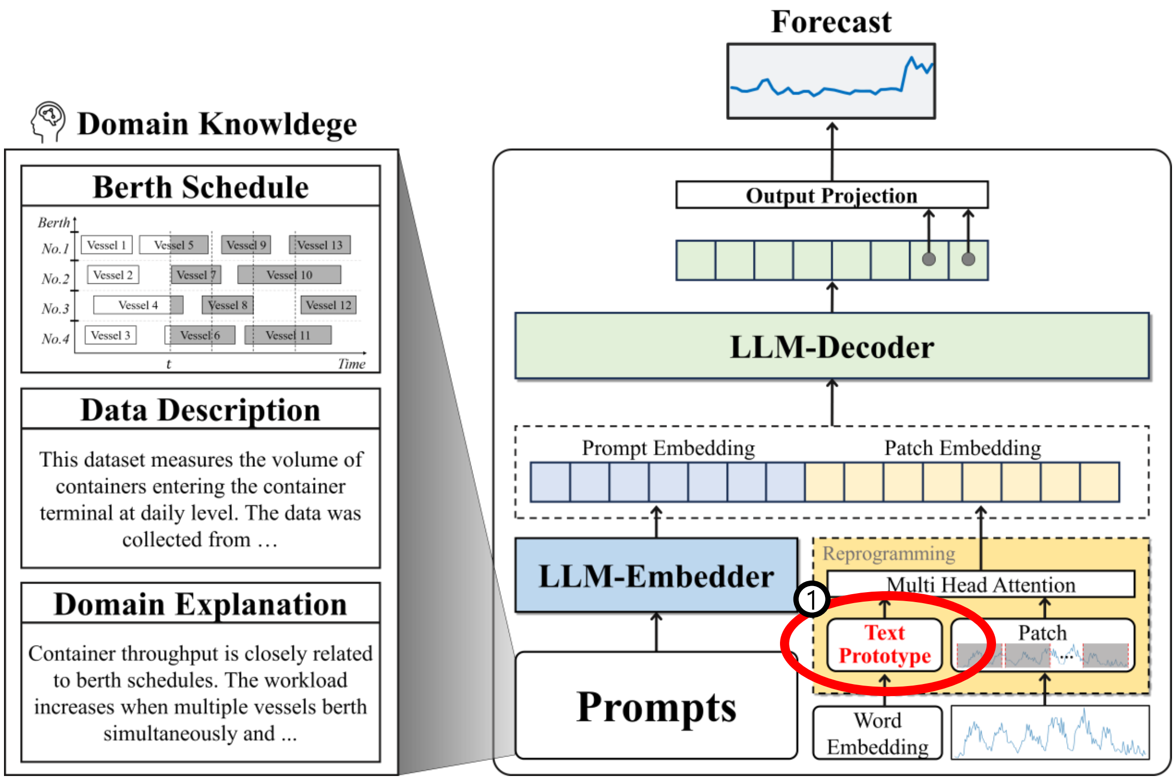
Table 2. Performance Comparison with Domain Knowledge Prompts

Parameters	w/o Prompt		w/ Prompt		w/ Prompt+Domain	
	MSE	MAE	MSE	MAE	MSE	MAE
0.5B	0.826	0.717	0.766	0.680	0.660	0.627
1.5B	1.031	0.825	0.918	0.761	0.620	0.614
3B	1.038	0.824	1.037	0.789	**0.560**	**0.573**
7B	**0.773**	**0.697**	**0.742**	**0.651**	0.724	0.670
14B	0.842	0.734	0.784	0.692	0.659	0.626
Avg	0.902	0.759	0.849	0.715	0.645	0.622
Std	0.111	0.054	0.112	0.051	0.054	0.031

3. 실제 항만 물류 분야의 적용 사례

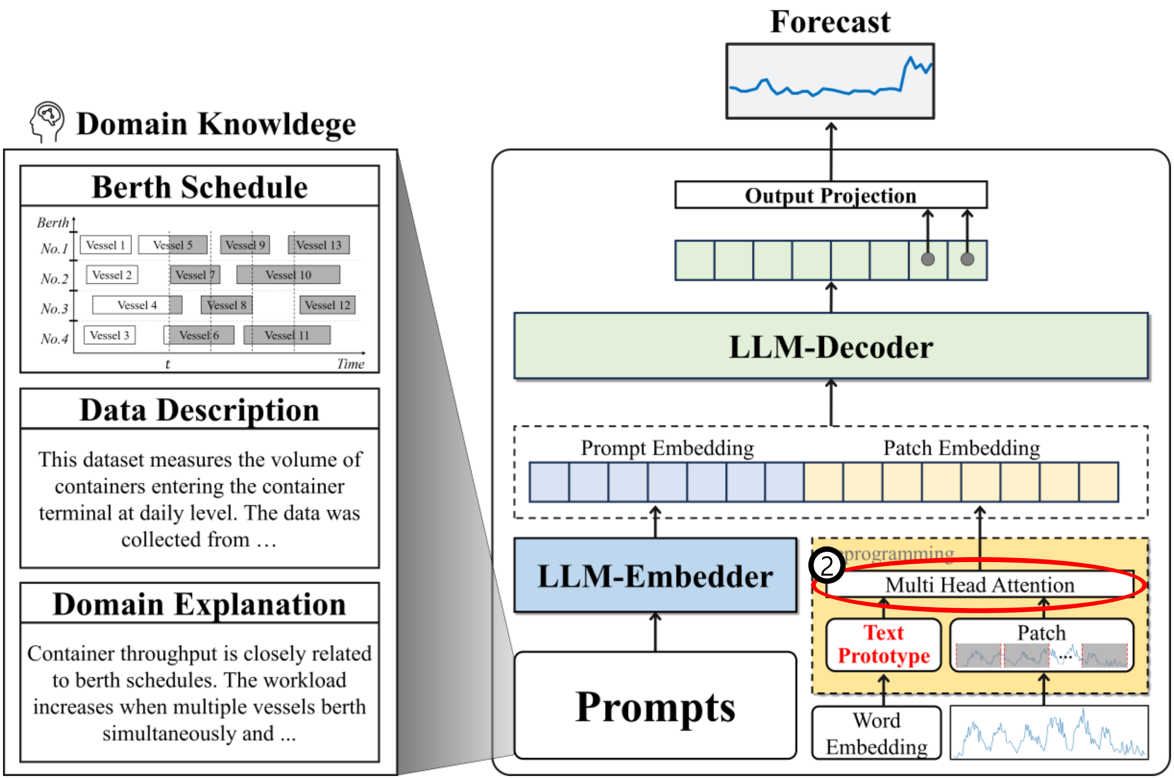
Interpretation Target 1

Word embedding에서 의미 있는 단어들을 추출하는가?



Interpretation Target 2

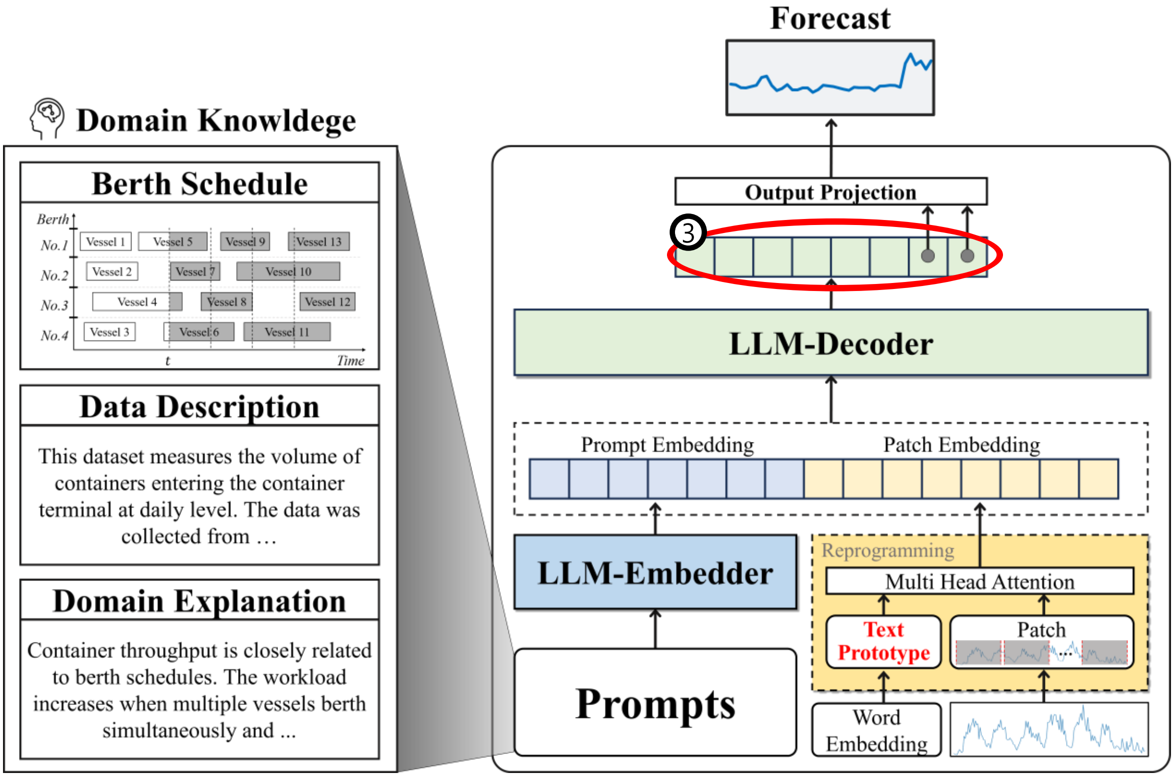
Text prototype과 time series data가 잘 결합하는가?



3. 실제 항만 물류 분야의 적용 사례

Interpretation Target 3

LLM-Decoder가 생성하는 embedding은 어떤 정보들을 담고 있는가?

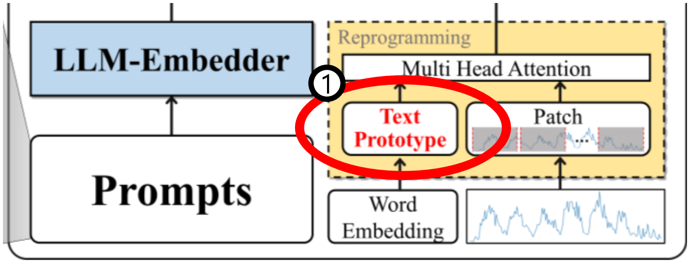


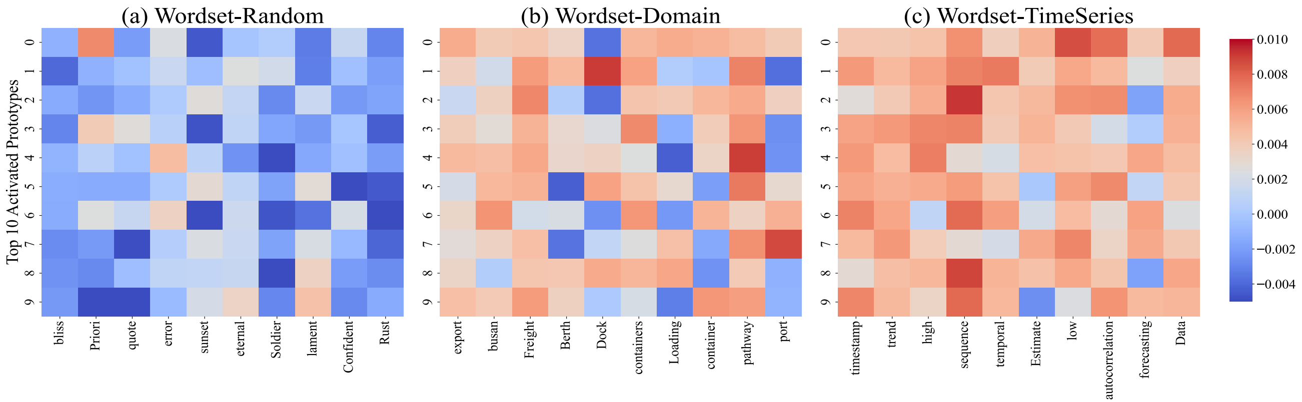
3. 실제 항만 물류 분야의 적용 사례

Interpretation Target 1

Word embedding에서 의미 있는 단어들을 추출하는가?

- word embedding에서 text prototype을 구성할 때 어떤 단어들이 가중치 활성도가 높은지 분석



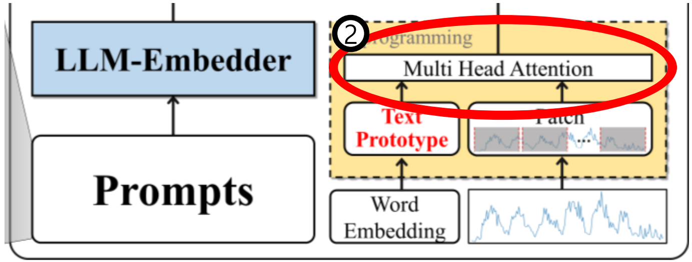


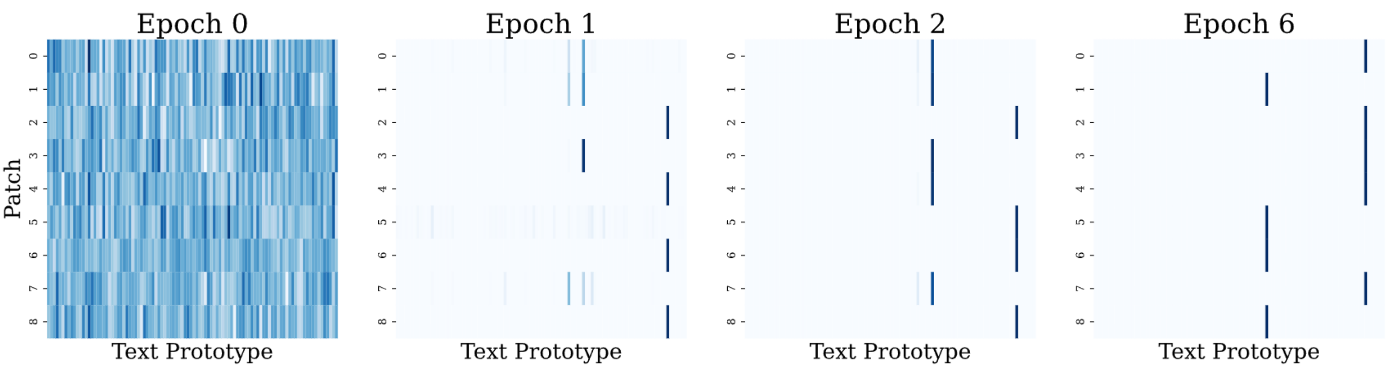
3. 실제 항만 물류 분야의 적용 사례

Interpretation Target 2

Text prototype과 time series data가 잘 결합하는가?

- Text Prototype과 time series data(patch)가 cross attention을 통해 결합하는 과정의 weight 시각화





Interpretation Target 3

LLM-Decoder가 생성하는 embedding은 어떤 정보들을 담고 있는가?

1) LLM(Decoder)이 생성하는 Token 정보를 Pre-Trained LLM에 입력

2) Pre-Trained LLM이 처음으로 생성하는 Token의 확률 분포에서 가장 높은 확률값을 가지는 단어들을 분석

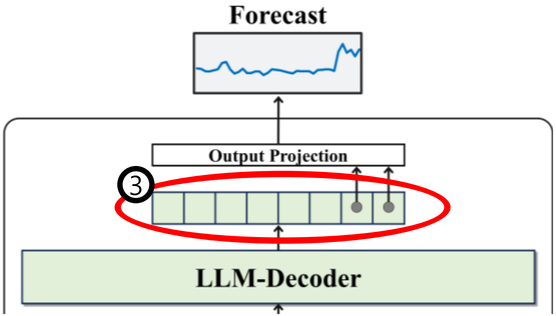


Table 3. High-Probability Words Extracted from Generated Token

Parameters	Semantic words
w/ Domain Knowledge	more, **information**, this, your, some, **future**, **holiday**, two, **public**, **additional**, new, any, **input**, **domain**, **context**, **detailed**, **season**, constraints, another, **seasonal**, **forecast**, **berth**, **details**, extra, **relevant**, comments, **data**, **up**, **specific**, code
w/o Domain Knowledge	cells, data, protein, water, cell, blood, components, proteins, people, energy, material, individuals, substances, elements

green: words related to time series data
red: words related to domain knowledge

향후 연구 방향

4. 향후 연구 방향

- LLM의 Application이 활발한 분야들(medical, law, finance, healthcare, etc)
    - Text를 이해하는데 특화되어 있는 정제된 텍스트(임상 기록, 법전과 재판 기록, 경제 뉴스 등)들이 많은 분야들
    - 논리적, 정확한 문법, 오타가 없음, 의미가 명확함, 통일된 용어를 사용

Article | Open access | Published: 07 June 2023

Health system-scale language models are all-purpose prediction engines

Lavender Yao Jiang, Xujin Chris Liu, Nima Pour Nejatian, Mustafa Nasir-Moin, Duo Wang, Anas Abidin, Kevin Eaton, Howard Antony Riina, Ilya Laufer, Paawan Punjabi, Madeline Miceli, Nora C. Kim, Cordelia Orillac, Zane Schnurman, Christopher Livia, Hannah Weiss, David Kurland, Sean Neifert, Yosef Dastagirzada, Douglas Kondziolka, Alexander T. M. Cheung, Grace Yang, Ming Cao, Mona Flores, ... Eric Karl Oermann

+ Show authors

Nature **619**, 357–362 (2023) | Cite this article

This article has been updated

Abstract

Physicians make critical time-constrained decisions every day. Clinical predictive models can

4. 향후 연구 방향

- **Port Logistics**
  - 정제된 Text들이 없고, 항만에서 발생하는 수많은 event들을 기록하기 위해 Table 형태(Event Log)로 관리되는 데이터들이 대부분임
    - Ex) QC, TC, YT의 작업 이력, 컨테이너들의 게이트 In/Out 이력, 선박들의 입항 예정 정보, 컨테이너의 다양한 특징 등
  - 직접적인 관계를 찾고 이해하기에는 제한이 있지만 이들 간에는 분명한 관계가 존재하고, 현장의 관리자들은 대부분 이러한 Tabular 데이터들을 활용하여 의사결정을 내리고 있음

- **Future Works**
  - (Task Agnostic) Tabular Data들로부터 지식을 학습하고, 이들을 통해 여러 downstream task를 수행하는 Foundation Model
    - 앞서 언급한 Nature 논문과 비슷한 연구 방향이지만, 활용 가능한 데이터가 text 형태의 clinical note가 아닌 tabular data라는 것이 차이점

  - (Task Specific) task가 주어졌을 때, Tabular Data들로부터 관련 있는 정보를 찾아내고 이를 결합할 수 있는 새로운 Framework
    - 현재의 접근 방식: 관련 정보를 직접 Prompting하고 있음

Q&A