

Open Seminar

LLMs for Time Series: TIME-LLM, TIME FFM



2025. 11. 14. 금

Data Mining & Quality Analytics Lab.

김수림



About Me



❖ 김수림 (Soorim Kim)

- 고려대학교 산업경영공학과 석사과정 (2025.09 ~ Present)
- Data Mining & Quality Analytics Lab. (김성범 교수님)

❖ Research Interest

- Federated Learning
- Time Series Analysis
- Large Language Models

❖ Contact

- srkim1@korea.ac.kr

Contents

❖ Introduction

❖ Methods

- TIME-LLM: Time series forecasting by reprogramming large language models
- TIME-FFM: Towards LM-empowered federated foundation model for time series forecasting

❖ Conclusion

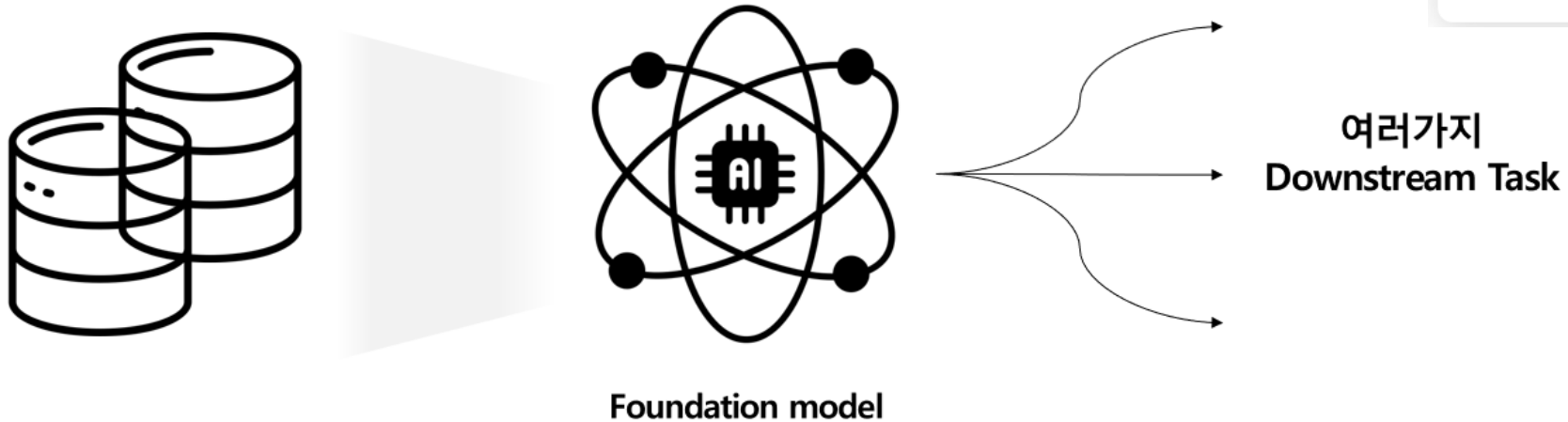
Introduction

❖ 왜 시계열 데이터는 Large Model이 없을까?

- 자연어 처리(NLP) 및 컴퓨터 비전(CV)에서는 사전 학습된 대규모 파운데이션 모델로 큰 발전을 이룸
- 하지만 시계열 데이터는 사전 학습을 하는 것이 어려움

What is a foundation model?

방대한 양의 데이터로 사전학습된 범용적인 모델




종료

DINOV2, DINOv3 : Self-supervised Foundation Model

DMQA Open Seminar

2025.10.31
Data Mining & Quality Analytics Lab.
0942

DINOv2, DINOv3 : Self-supervised Vision I

발표자:  이혜승

📅 2025년 10월 31일
🕒 오후 12시 ~
📺 온라인 비디오 시청 (YouTube)

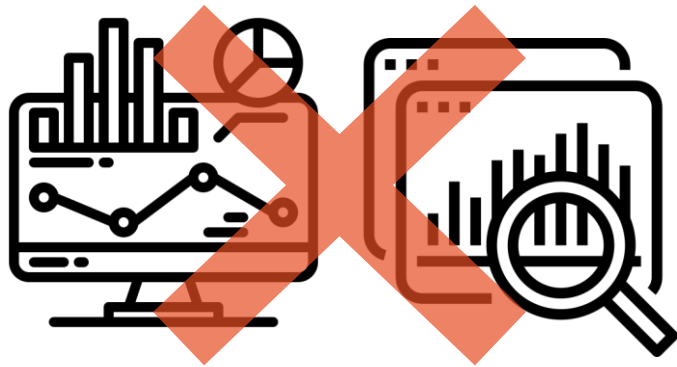
세미나 정보 보기 →

Introduction

❖ 왜 시계열 데이터는 Large Model이 없을까?

- 자연어 처리(NLP) 및 컴퓨터 비전(CV)에서는 사전 학습된 대규모 파운데이션 모델로 큰 발전을 이룸
- 하지만 시계열 데이터는 사전 학습을 하는 것이 어려움

파운데이션 모델을 학습시킬 대량의 데이터 부족



시계열 분석을 위한 가장 큰 데이터셋은 10GB 미만

Introduction

❖ 왜 시계열 데이터는 Large Model이 없을까?

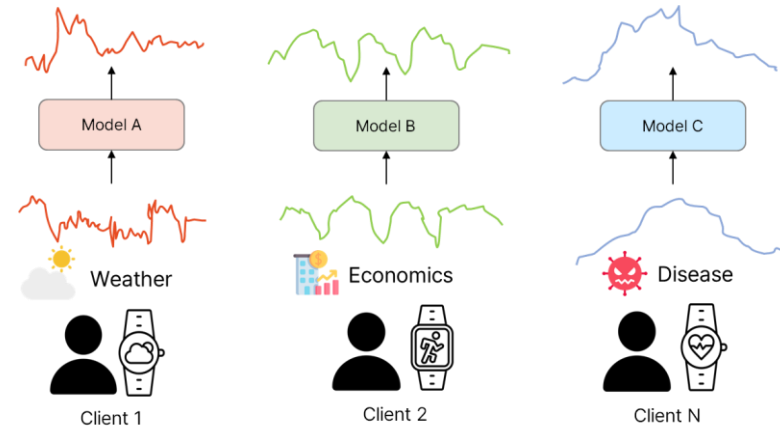
- 자연어 처리(NLP) 및 컴퓨터 비전(CV)에서는 사전 학습된 대규모 파운데이션 모델로 큰 발전을 이룸
- 하지만 시계열 데이터는 사전 학습을 하는 것이 어려움

파운데이션 모델을 학습시킬 대량의 데이터 부족



시계열 분석을 위한 가장 큰 데이터셋은 10GB 미만

데이터셋마다 고유 특징 및 형태가 모두 다름



시간 축, 변수 개수, scale, 통계적 수치 등..
→ 이를 통합한 대규모 데이터셋 구축 및 학습 어려움

Introduction

❖ 왜 시계열 데이터는 Large Model이 없을까?

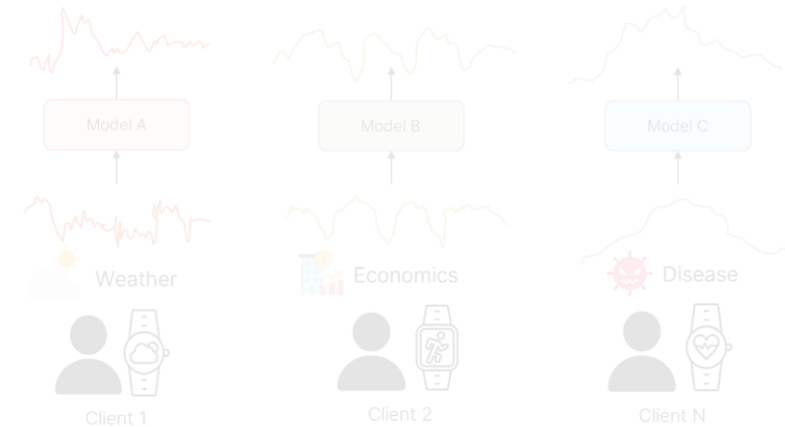
- 자연어 처리(NLP) 및 컴퓨터 비전(CV)에서는 사전 학습된 대규모 파운데이션 모델로 큰 발전을 이룸
- 하지만 시계열 데이터는 사전 학습을 하는 것이 어려움

언어 모델을 사용하자!

파운데이션 모델을 학습시킬 때와 다른 데이터셋과 고유 특징 및 형태가 모두 다름



시계열 분석을 위한 가장 큰 데이터셋은 10GB 미만



시간 축, 변수 개수, scale, 통계적 수치 등..
→ 이를 통합한 대규모 데이터셋 구축 및 학습 어려움

Introduction

❖ 왜 시계열 데이터는 Large Model이 없을까?

- 자연어 처리(NLP) 및 컴퓨터 비전(CV)에서는 사전 학습된 대규모 파운데이션 모델로 큰 발전을 이룸
- 하지만 시계열 데이터는 사전 학습을 하는 것이 어려움

언어 모델을 사용하자! 왜?

파운데이션 모델을 학습시킬 때 사용하는 데이터의 종류와 형태가 모두 다름

시계열과 언어 모두 시간과 순서에 따라 의미가 변하는 시퀀스 데이터



언어 모델이 next text token을 예측하는 것처럼 시계열의 next timestep을 예측할 수 있을 것이다

→ 이를 통합한 대규모 데이터셋 구축 및 학습 어려움

Introduction

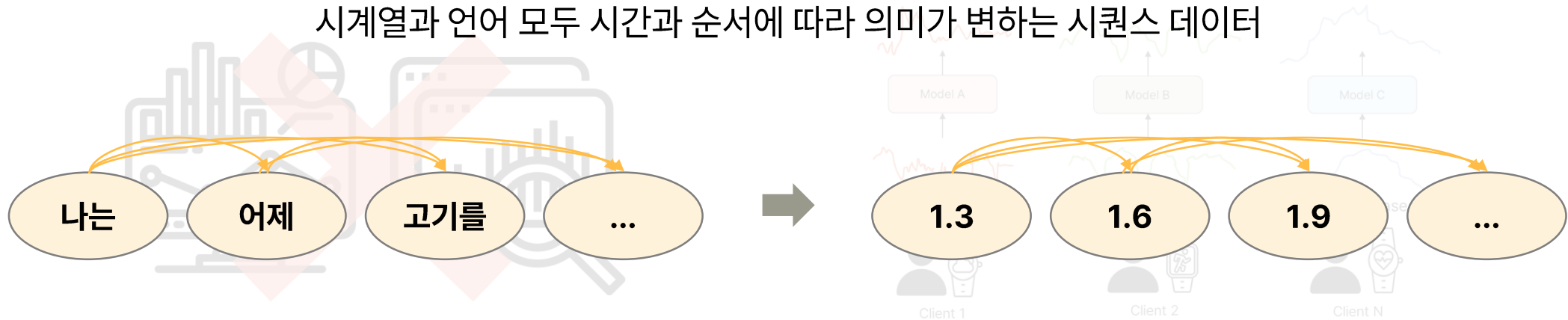
❖ 왜 시계열 데이터는 Large Model이 없을까?

- 자연어 처리(NLP) 및 컴퓨터 비전(CV)에서는 사전 학습된 대규모 파운데이션 모델로 큰 발전을 이룸
- 하지만 시계열 데이터는 사전 학습을 하는 것이 어려움

언어 모델을 사용하자! 왜?

파운데이션 모델을 학습시킬 때, 데이터의 종류와 형태가 모두 다름

시계열과 언어 모두 시간과 순서에 따라 의미가 변하는 시퀀스 데이터

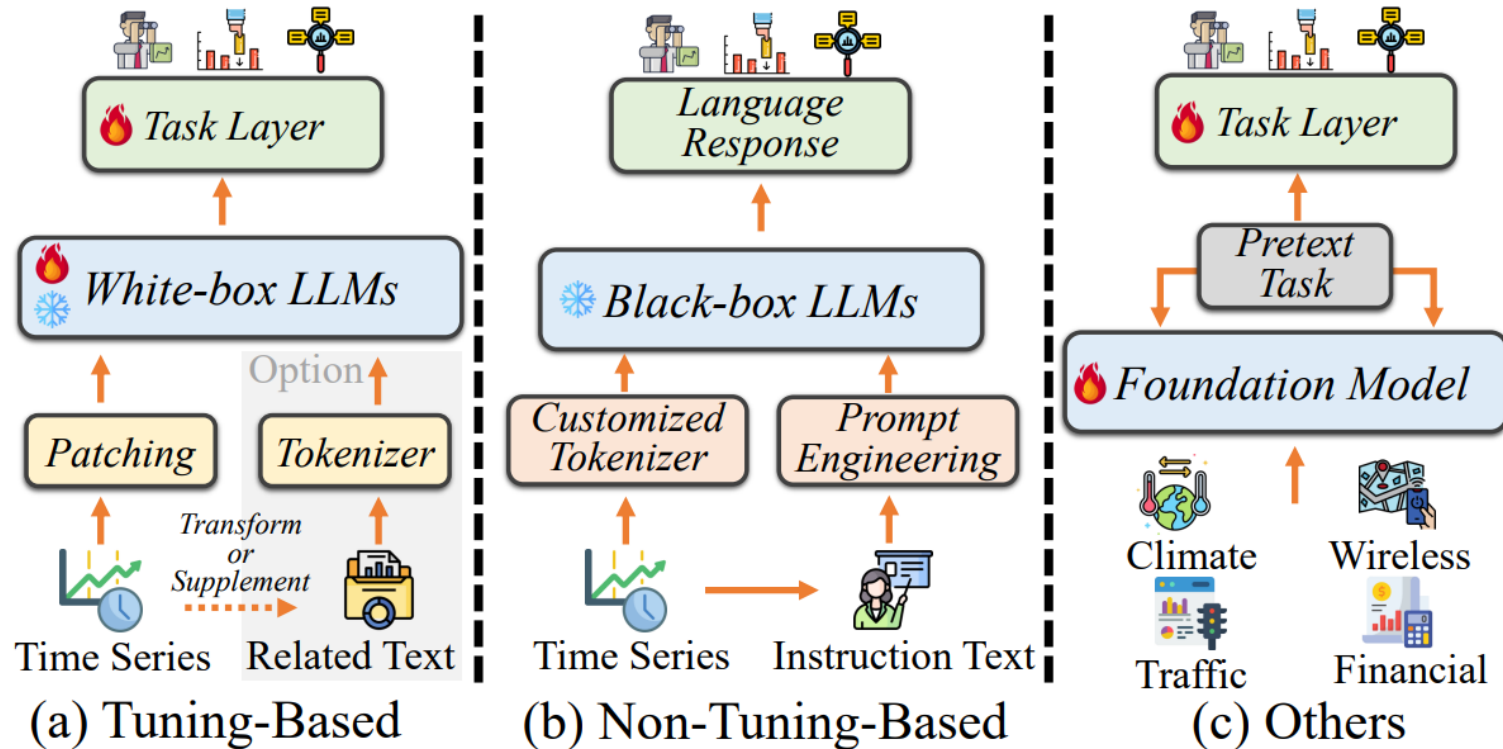


언어 모델이 순차적 의존 관계를 충분히 학습하였다면, 시계열 데이터 내의 패턴도 잘 포착할 것이다
→ 이를 통합한 대규모 데이터셋 구축 및 학습 어려움

Introduction

❖ LLM-centered predictor

- LLM 중심의 시계열 예측 모델



TIME-LLM

❖ TIME-LLM: Time Series Forecasting by Reprogramming Large Language Models (ICRL, 2024)

- 사전 학습된 LLM을 활용하여 시계열 예측 파운데이션 모델 개발

📖 TIME-LLM: TIME SERIES FORECASTING BY REPROGRAMMING LARGE LANGUAGE MODELS

Ming Jin^{1*}, Shiyu Wang^{2*}, Lintao Ma², Zhixuan Chu², James Y. Zhang², Xiaoming Shi², Pin-Yu Chen³, Yuxuan Liang⁶, Yuan-Fang Li¹, Shirui Pan⁴, Qingsong Wen⁵†

¹Monash University ²Ant Group ³IBM Research ⁴Griffith University ⁵Alibaba Group

⁶The Hong Kong University of Science and Technology (Guangzhou)

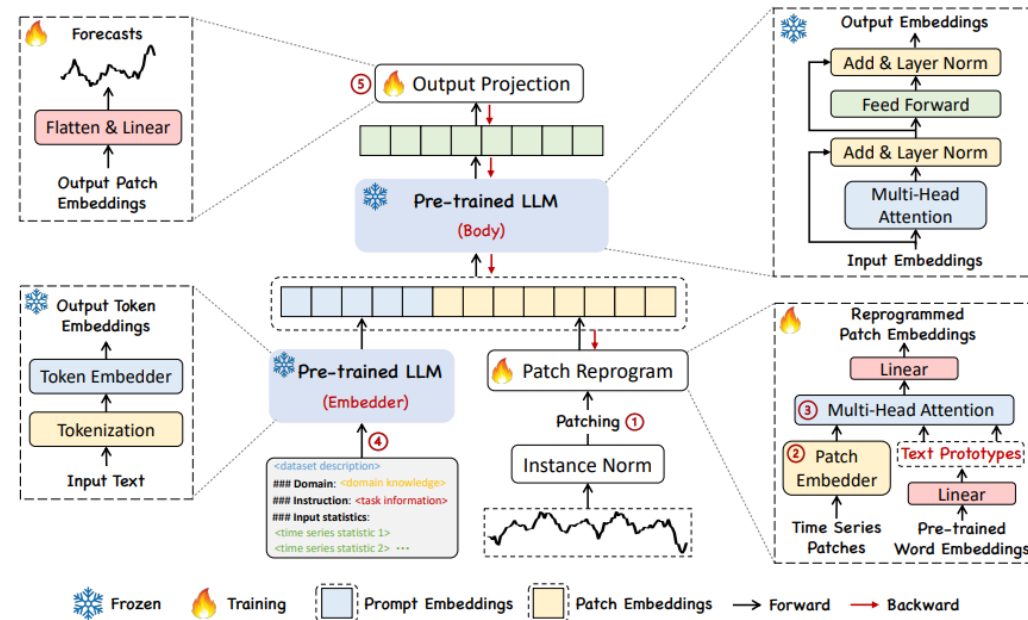
{ming.jin, yuanfang.li}@monash.edu, pin-yu.chen@ibm.com

yuxliang@outlook.com, s.pan@griffith.edu.au, qingsongedu@gmail.com

{weiming.wsy, lintao.mlt, chuzhixuan.czx, james.z, peter.sxm}@antgroup.com

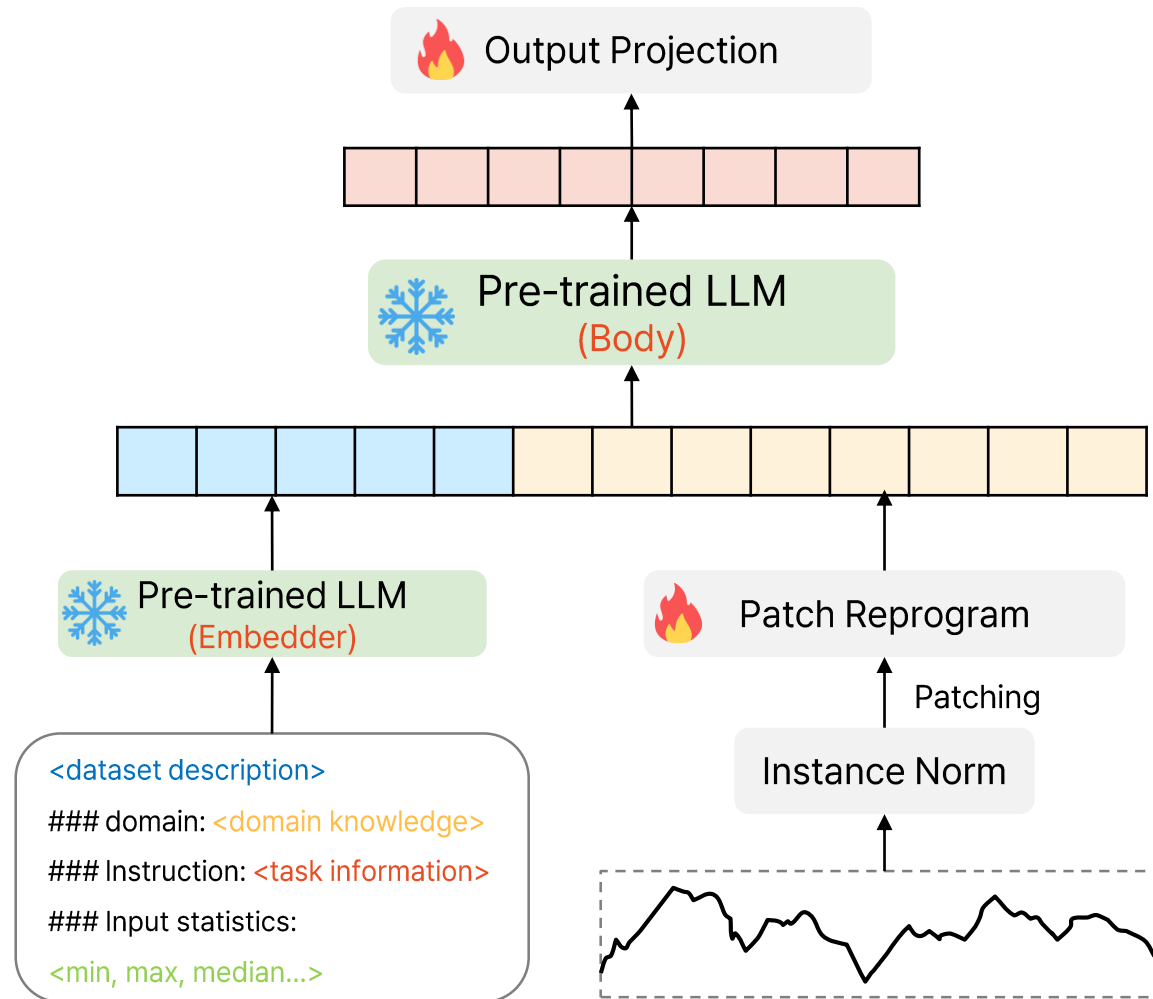
ABSTRACT

Time series forecasting holds significant importance in many real-world dynamic systems and has been extensively studied. Unlike natural language process (NLP) and computer vision (CV), where a single large model can tackle multiple tasks, models for time series forecasting are often specialized, necessitating distinct designs for different tasks and applications. While pre-trained foundation models have made impressive strides in NLP and CV, their development in time series domains has been constrained by data sparsity. Recent studies have revealed that large language models (LLMs) possess robust pattern recognition and reasoning abilities over complex sequences of tokens. However, the challenge remains in effectively aligning the modalities of time series data and natural language to leverage these capabilities. In this work, we present TIME-LLM, a reprogramming framework to repurpose LLMs for general time series forecasting with the backbone language models kept intact. We begin by reprogramming the input time series with text prototypes before feeding it into the frozen LLM to align the two modalities. To augment the LLM's ability to reason with time series data, we propose Prompt-as-Prefix (PaP), which enriches the input context and directs the transformation of reprogrammed input patches. The transformed time series patches from the LLM are finally projected to obtain the forecasts. Our comprehensive evaluations demonstrate that TIME-LLM is a powerful time series learner that outperforms state-of-the-art, specialized forecasting models. Moreover, TIME-LLM excels in both few-shot and zero-shot learning scenarios. The code is made available at <https://github.com/KimMeen/Time-LLM>



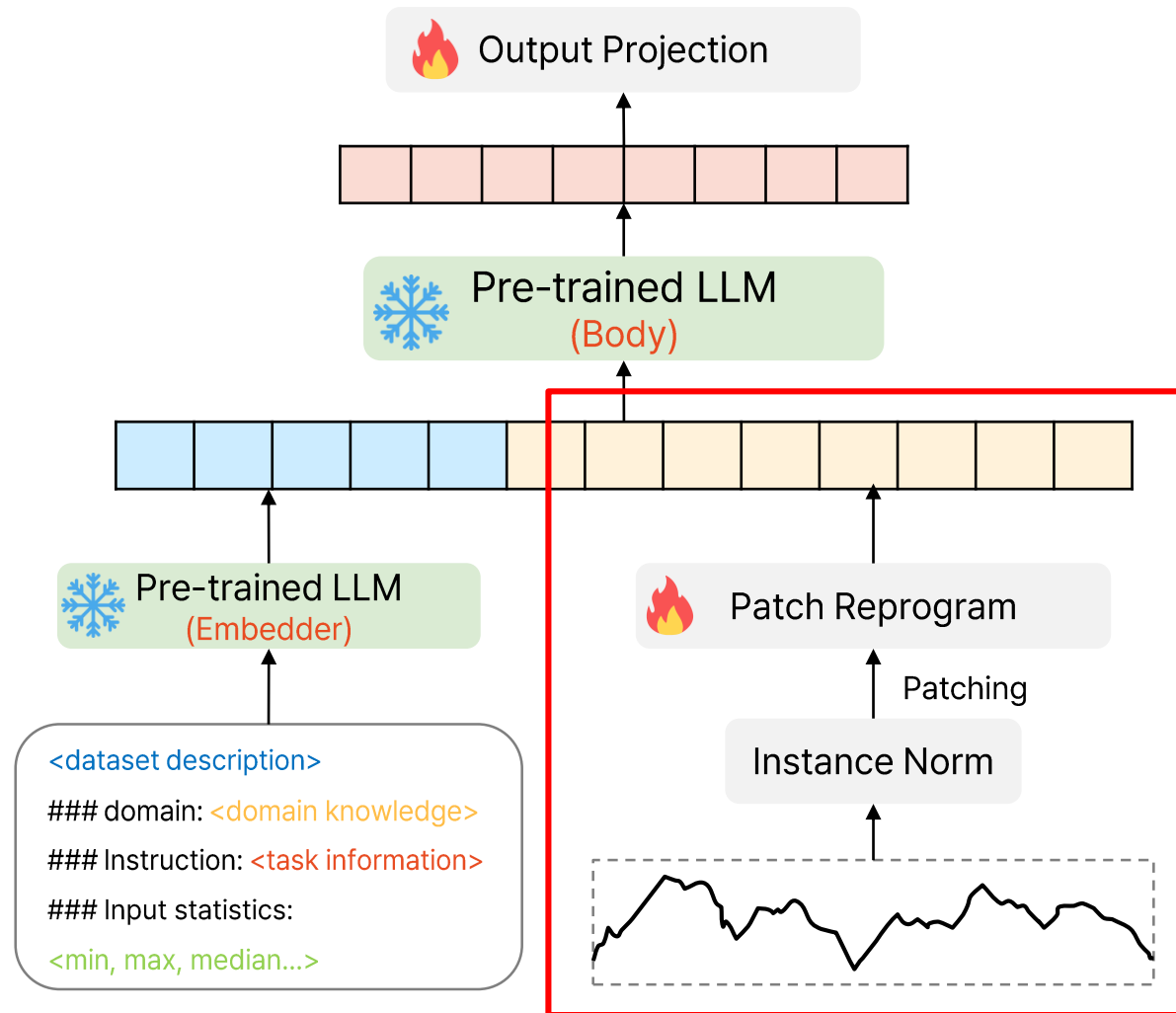
TIME-LLM

❖ Contribution



TIME-LLM

❖ Contribution



Patch Reprogramming

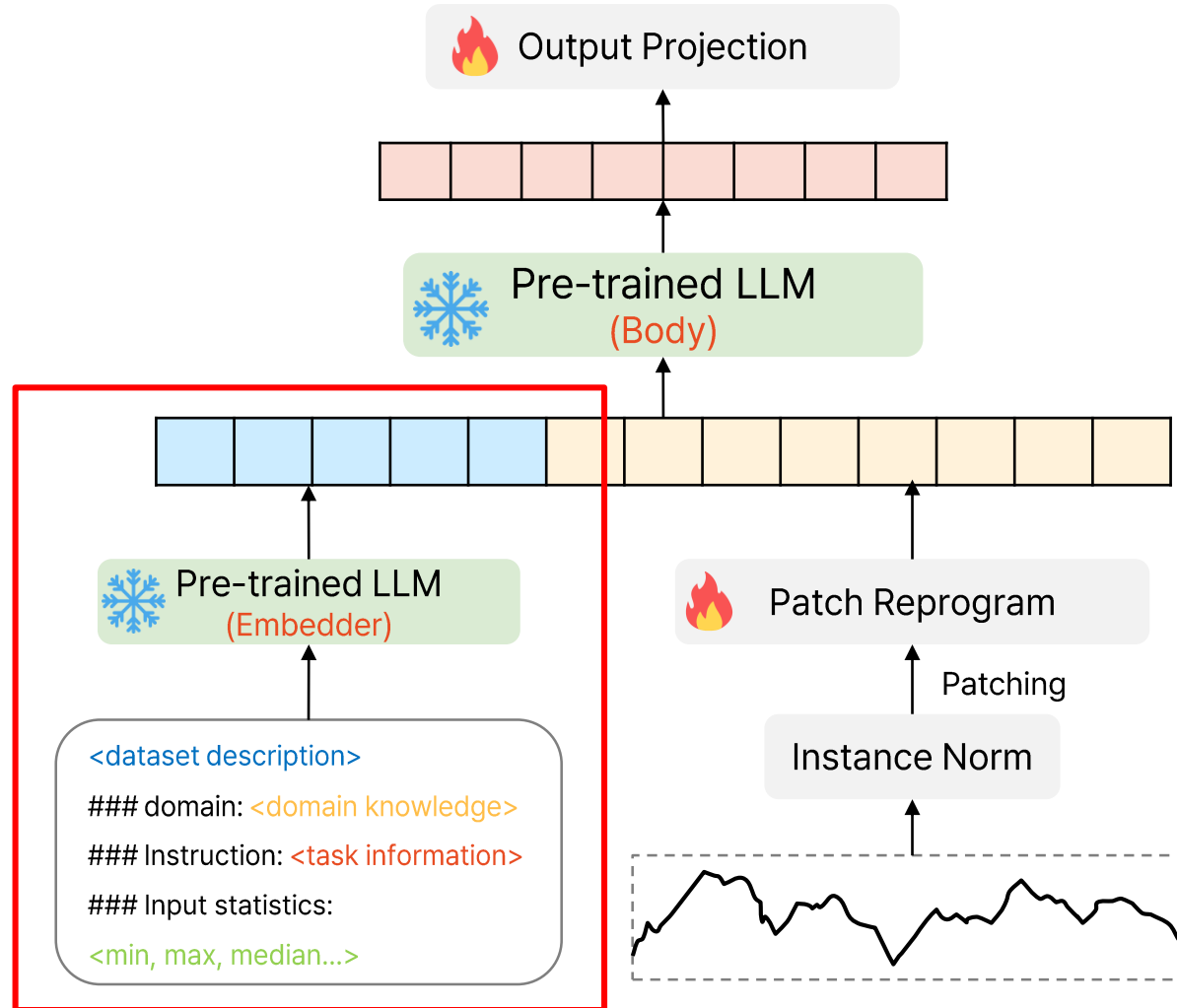
언어 모델 입력을 위한
모달리티 정렬
시계열 → 자연어

TIME-LLM

❖ Contribution

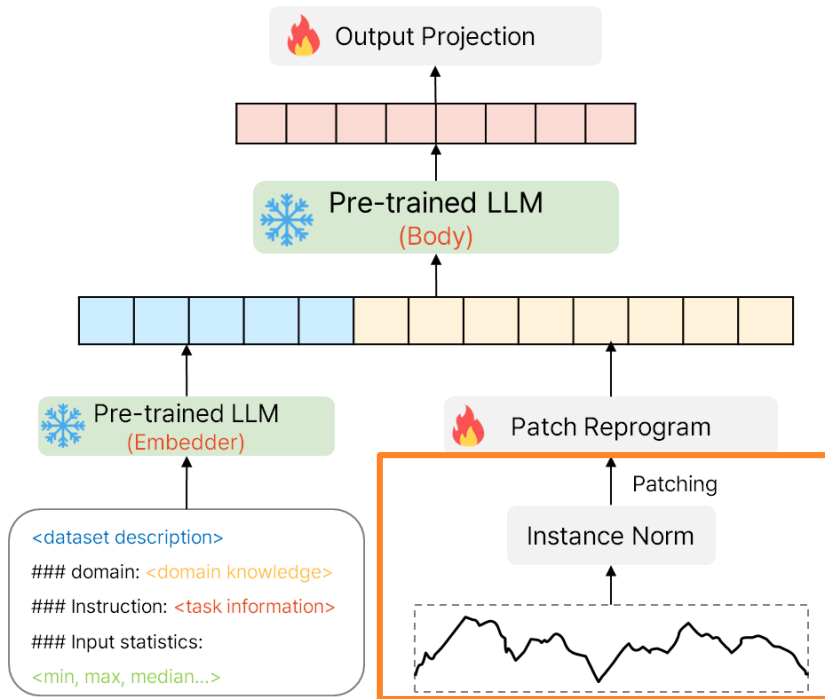
Prompt-as-Prefix

시계열 데이터의 추가적인 정보를
text 형태의 접두사로 입력



TIME-LLM

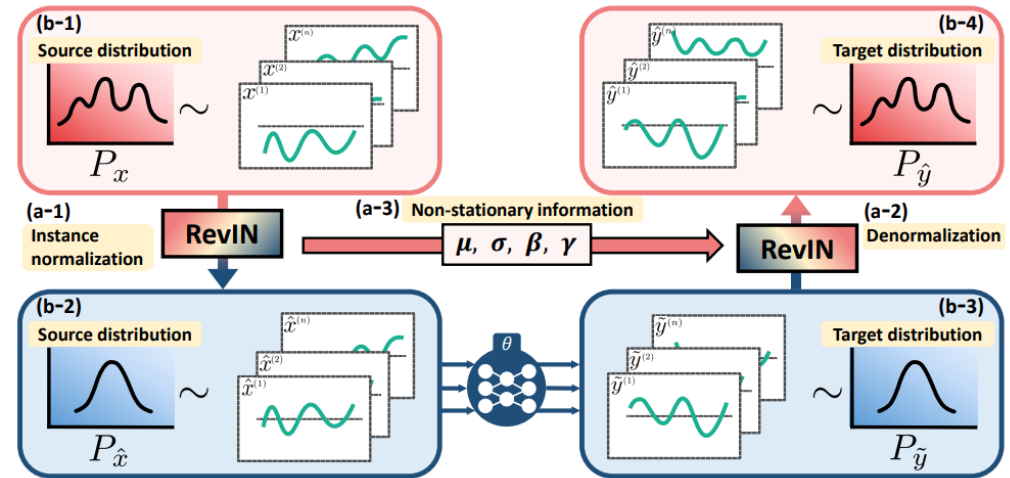
❖ Input Embedding



시계열 정규화 & Patching

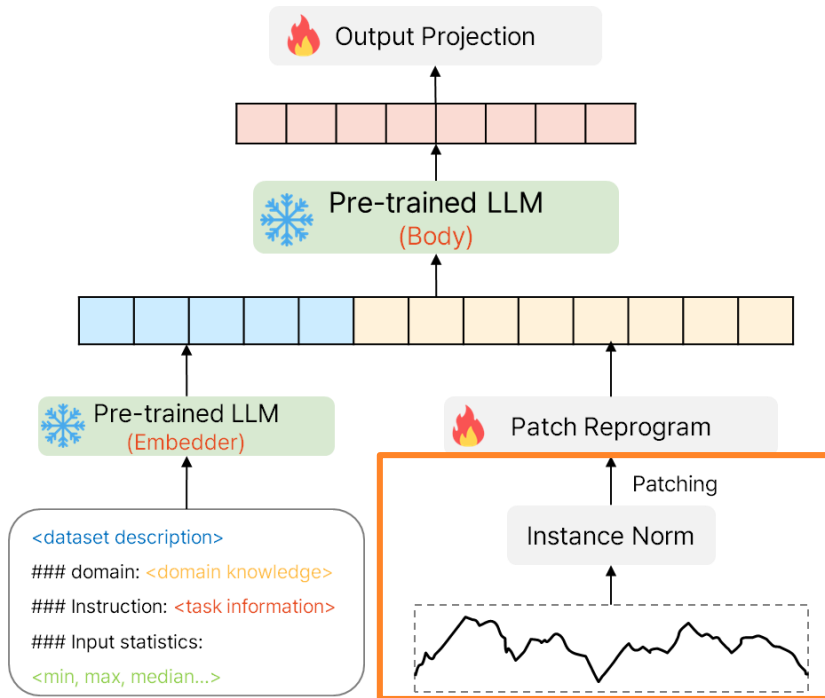
Reversible instance normalization(RevIN)

Learnable parameter를 두어
training data와 target data 사이의 불일치로
예측 모델의 성능이 낮아지는 것을 방지



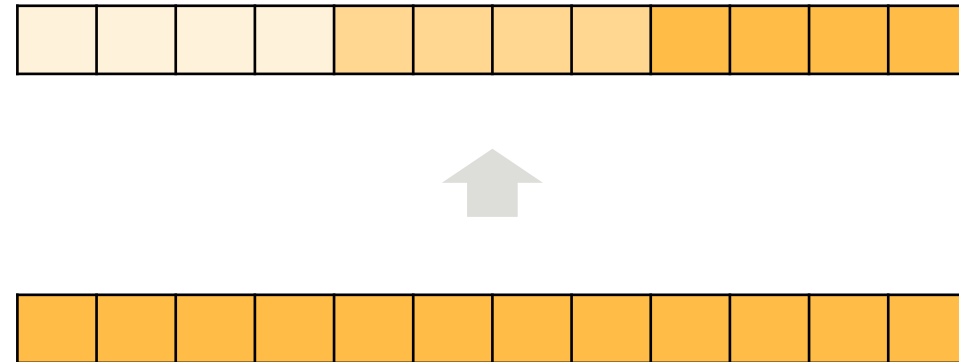
TIME-LLM

❖ Input Embedding



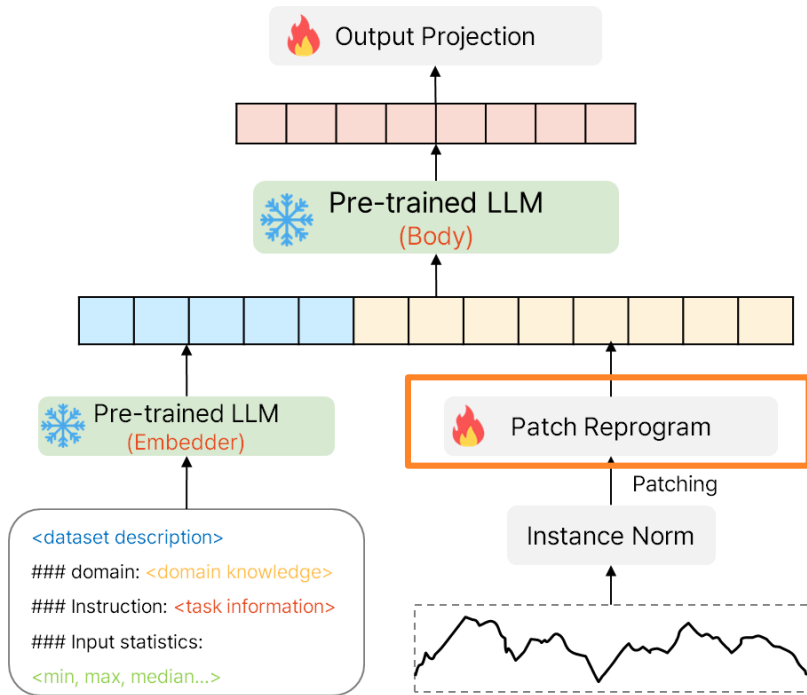
시계열 정규화 & Patching

단일 시점이 아닌 patch 단위로 처리하여
Local 정보 보존 및 자원 cost 감소



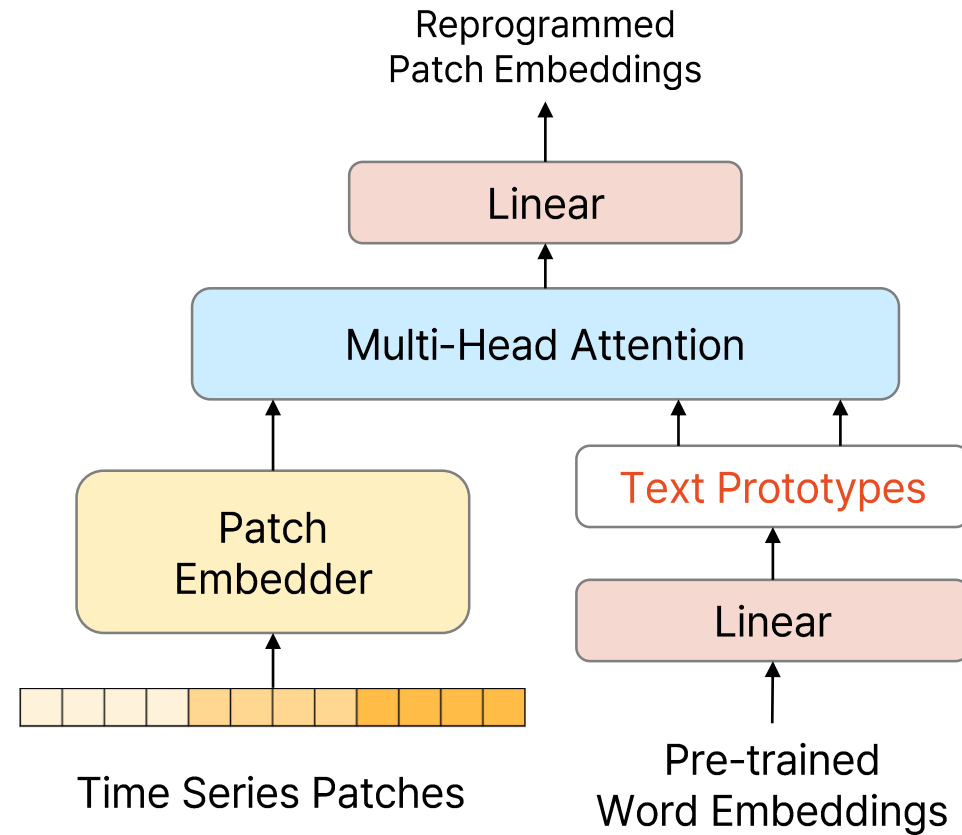
TIME-LLM

❖ Modality Alignment



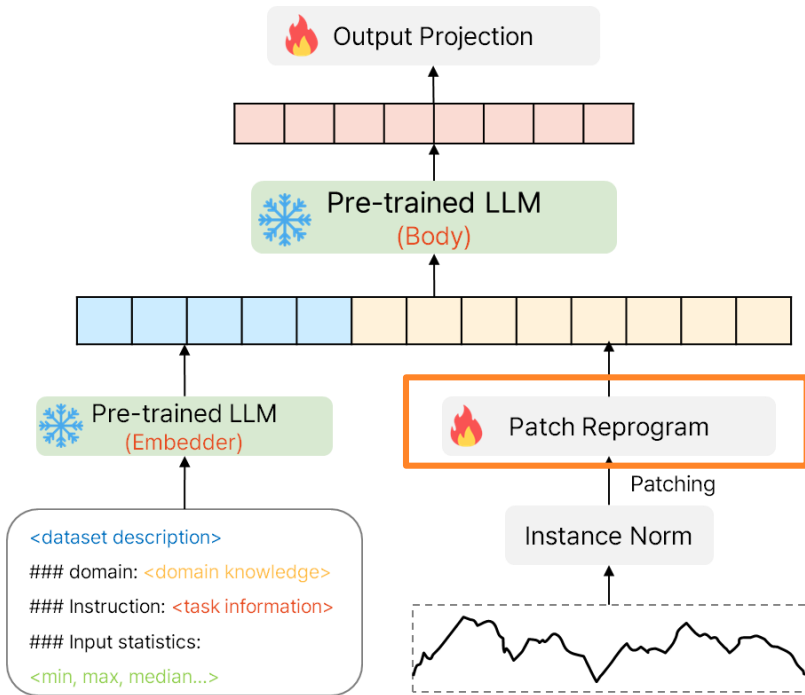
🔥 Patch Reprogram

언어 모델을 사용하기 위해 시계열 특징을 자연어 도메인에 정렬



TIME-LLM

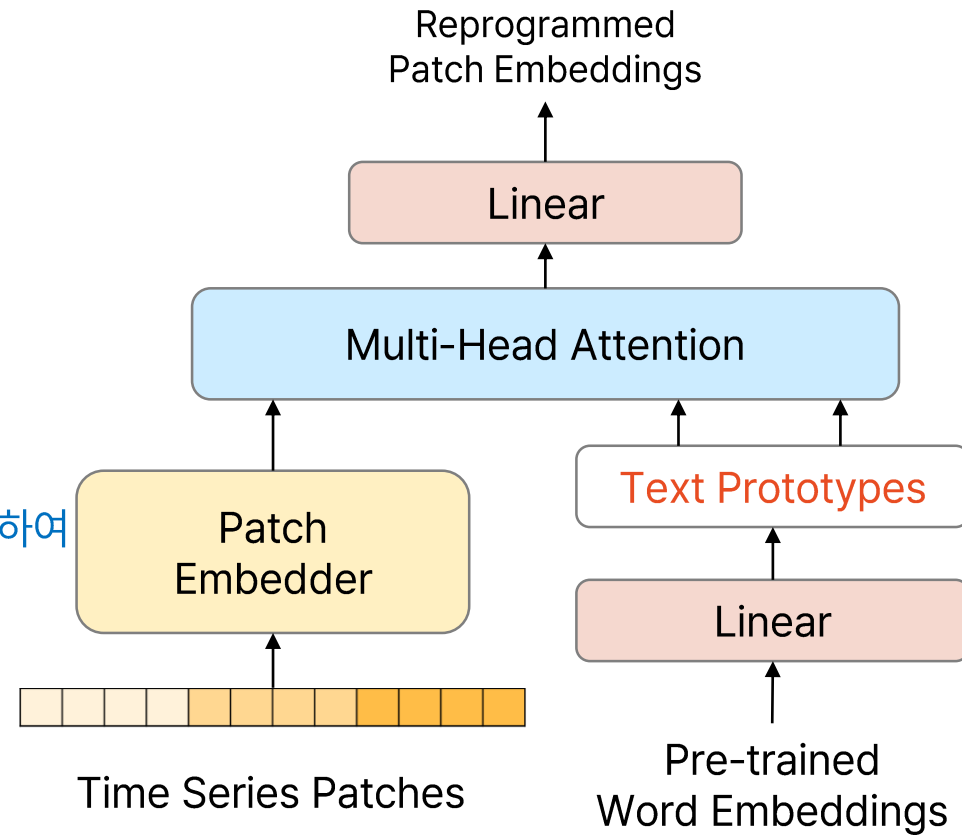
❖ Modality Alignment



Linear layer 통과하여 embedding

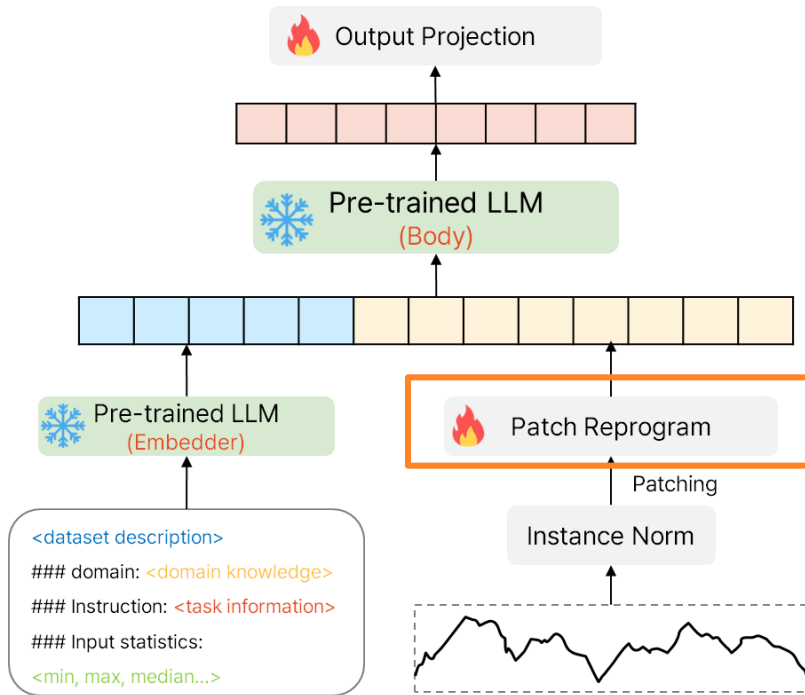
🔥 Patch Reprogram

언어 모델을 사용하기 위해 시계열 특징을 자연어 도메인에 정렬



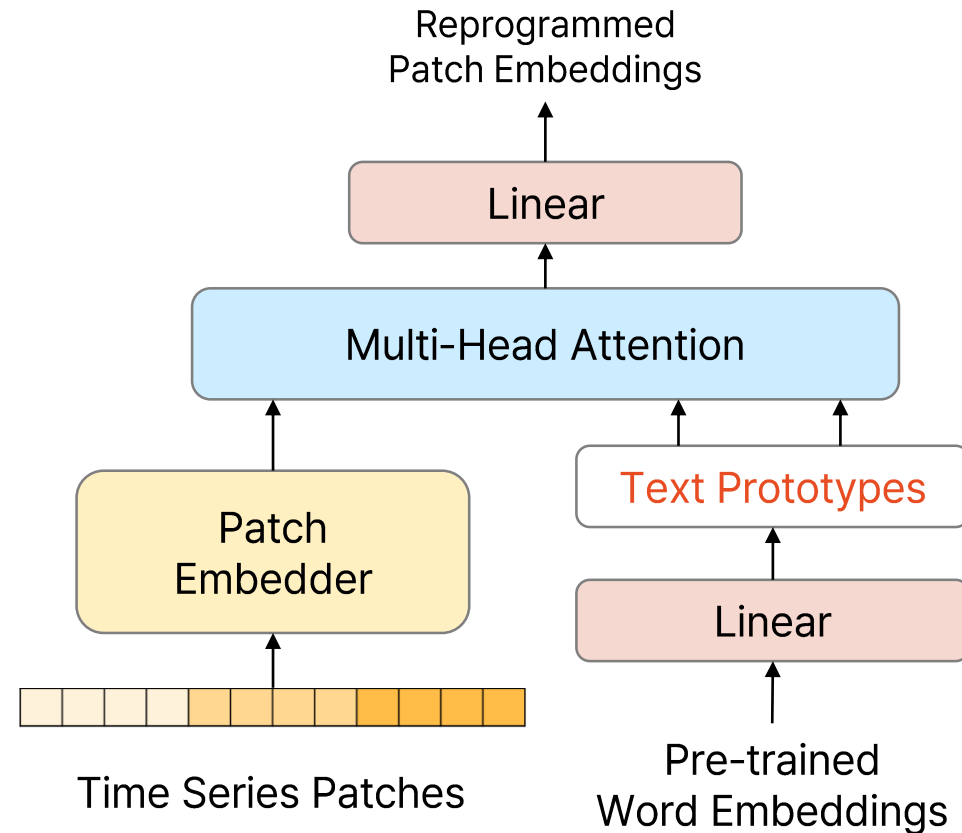
TIME-LLM

❖ Modality Alignment



🔥 Patch Reprogram

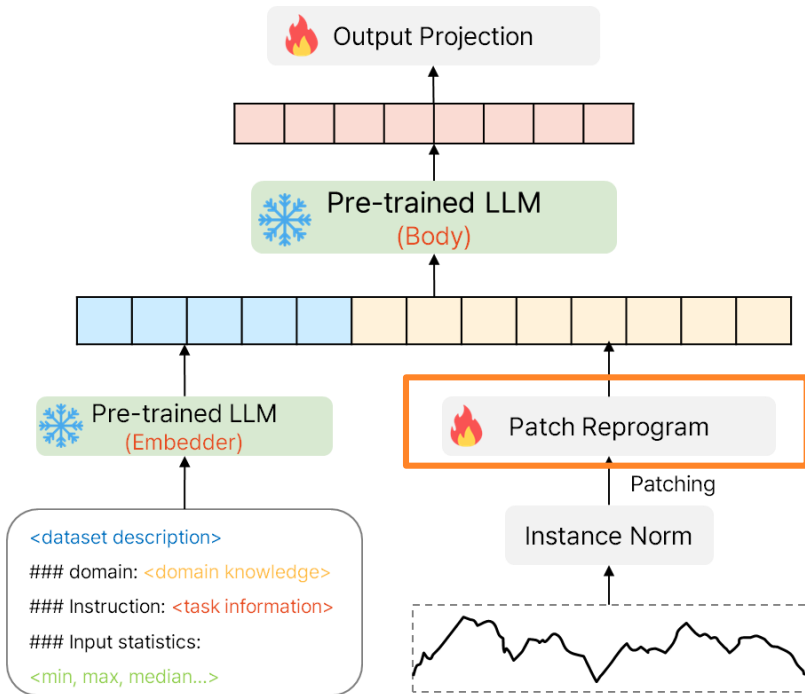
언어 모델을 사용하기 위해 시계열 특징을 자연어 도메인에 정렬



Linear layer를 통해
전체 Vocab 차원 축소
→ 핵심 단어만 남기기

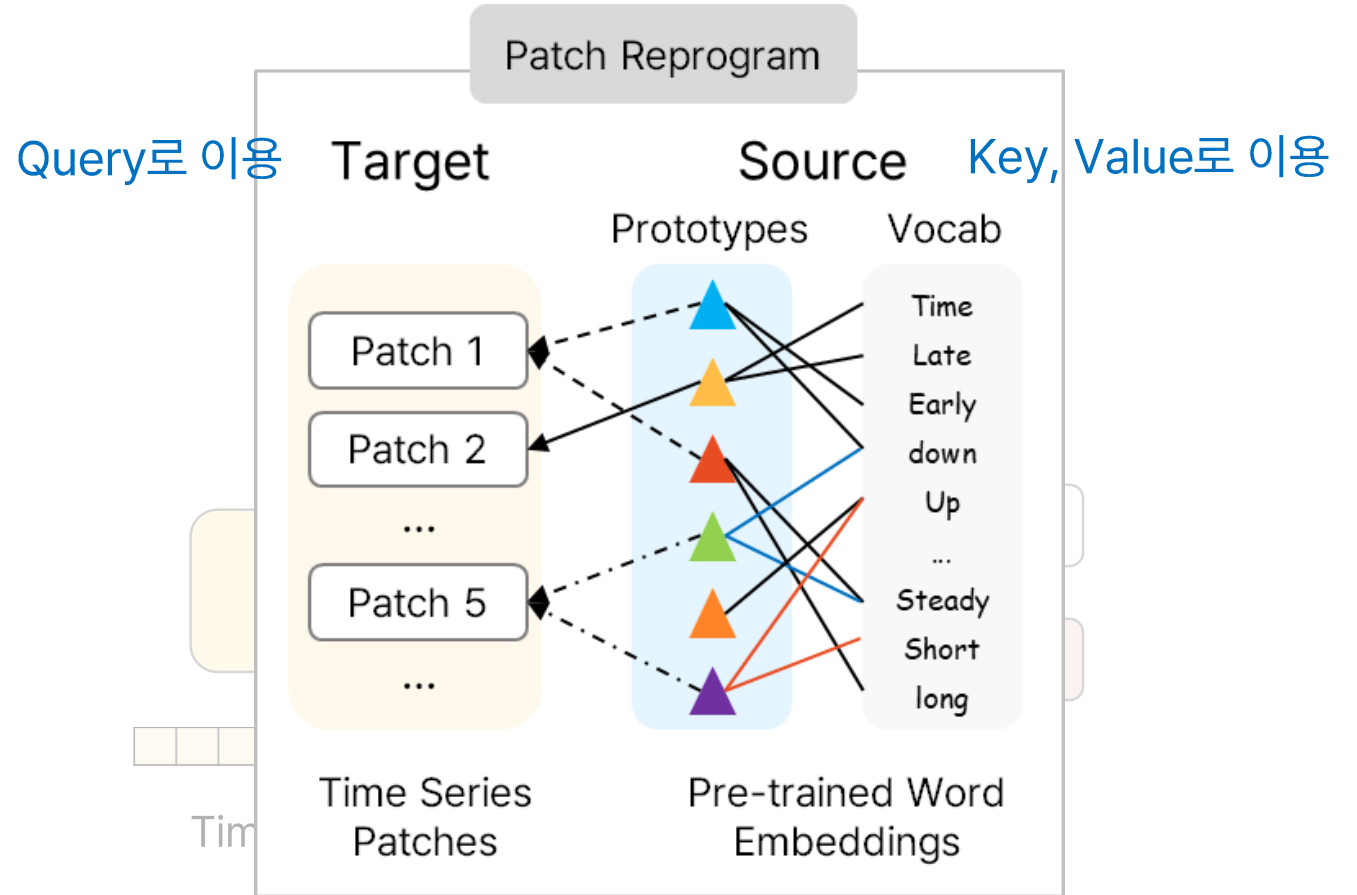
TIME-LLM

❖ Modality Alignment



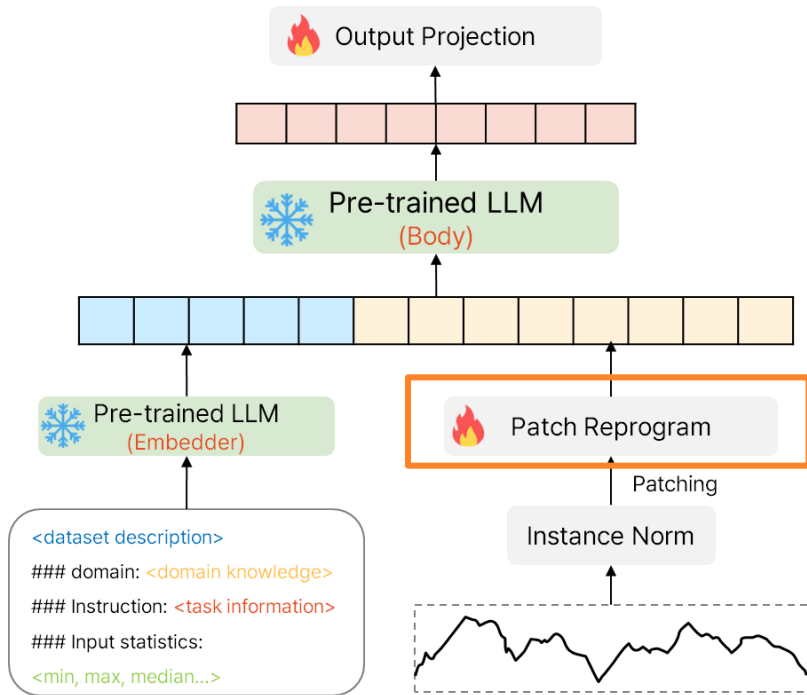
🔥 Patch Reprogram

언어 모델을 사용하기 위해 시계열 특징을 자연어 도메인에 정렬



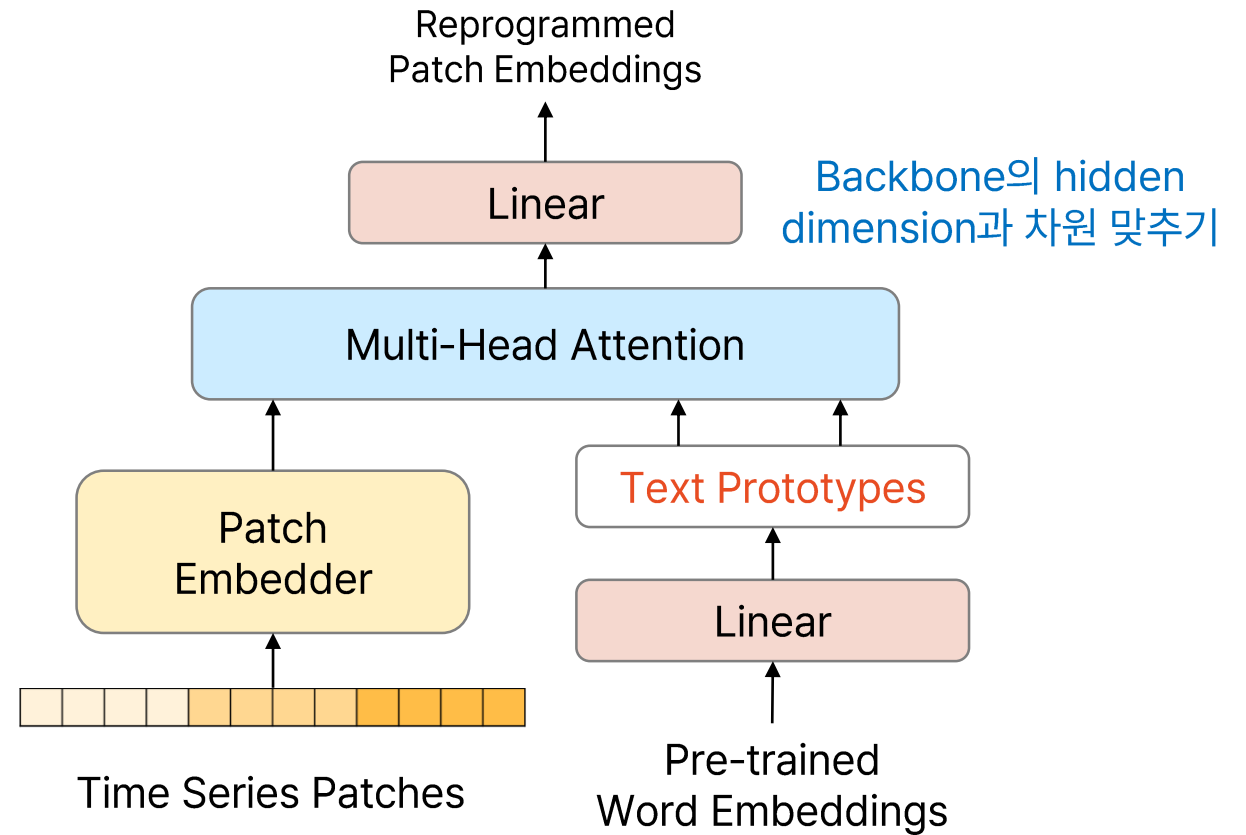
TIME-LLM

❖ Modality Alignment



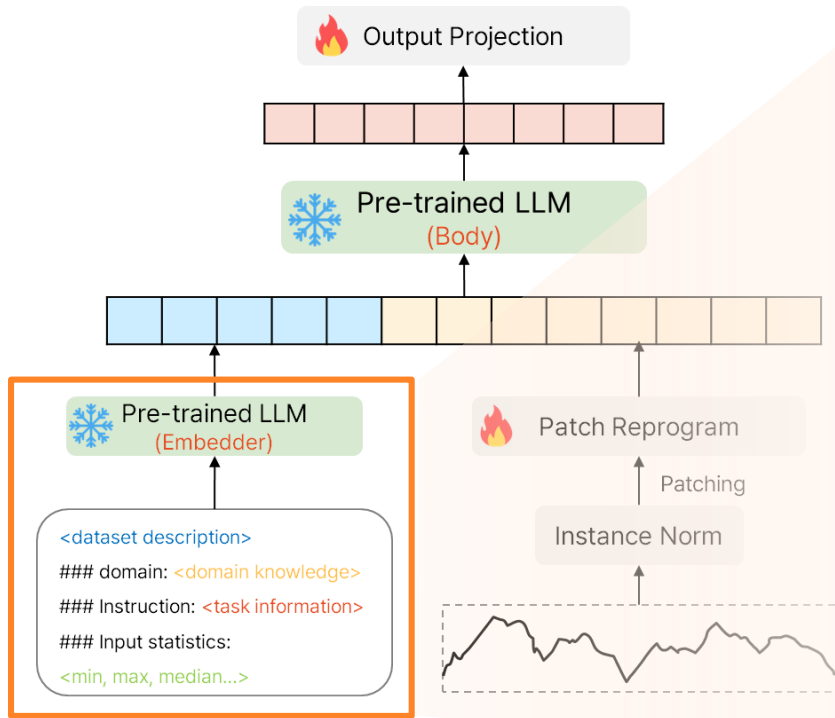
🔥 Patch Reprogram

언어 모델을 사용하기 위해 시계열 특징을 자연어 도메인에 정렬



TIME-LLM

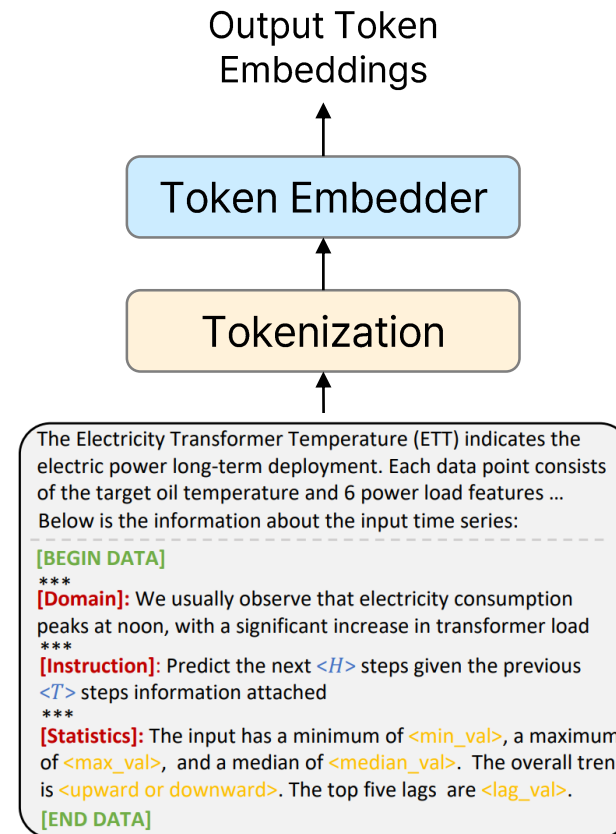
❖ Prompt-as-Prefix



데이터셋 정보
task 설명
통계 값 (단변량 기준)
...

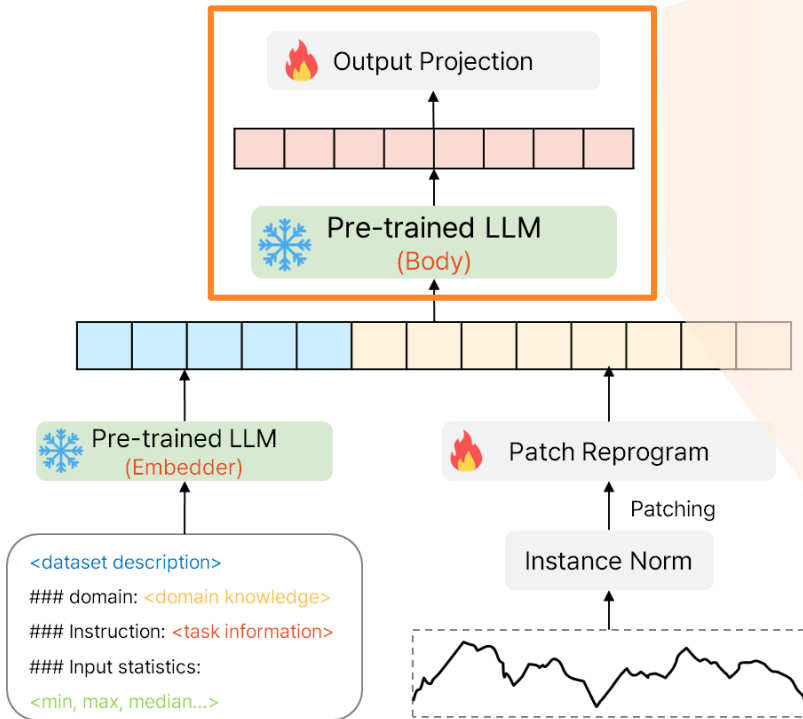
Prompt-as-Prefix

시계열의 정보를 text prompt로 구성하여 접두사로 사용
→ 시계열 patch를 더욱 잘 이해할 수 있음

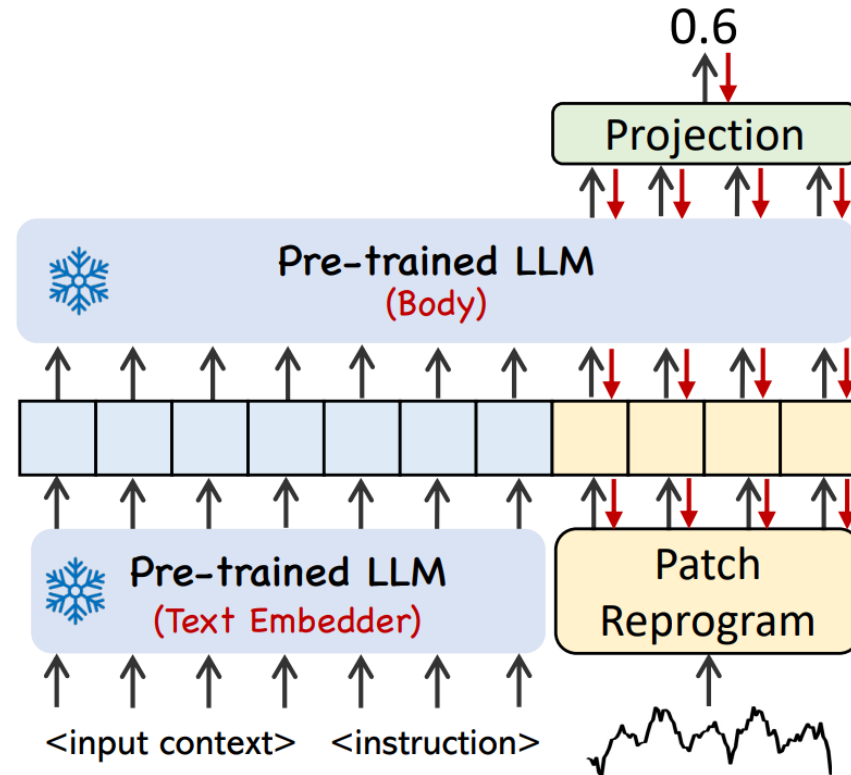


TIME-LLM

❖ Output Projection



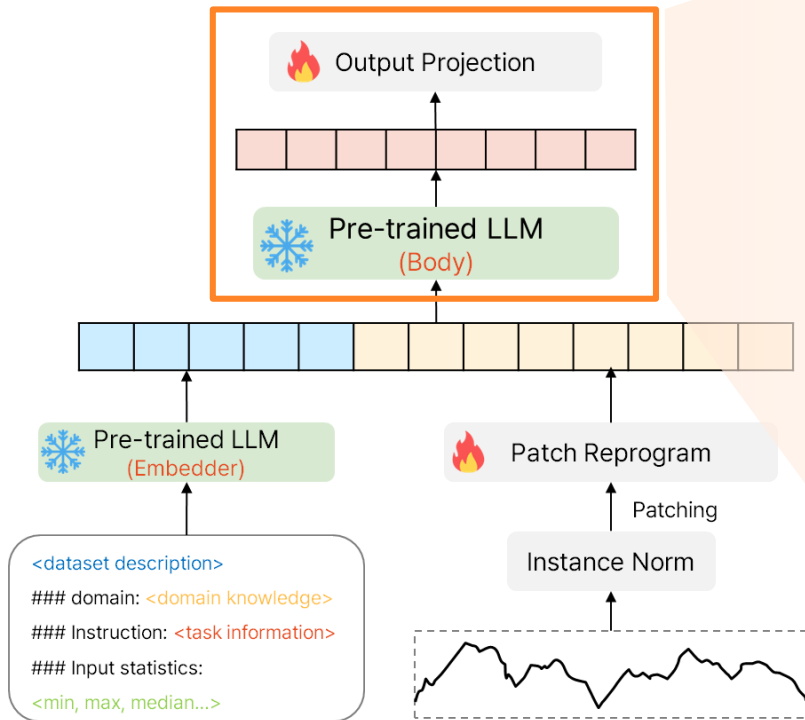
Output Projection



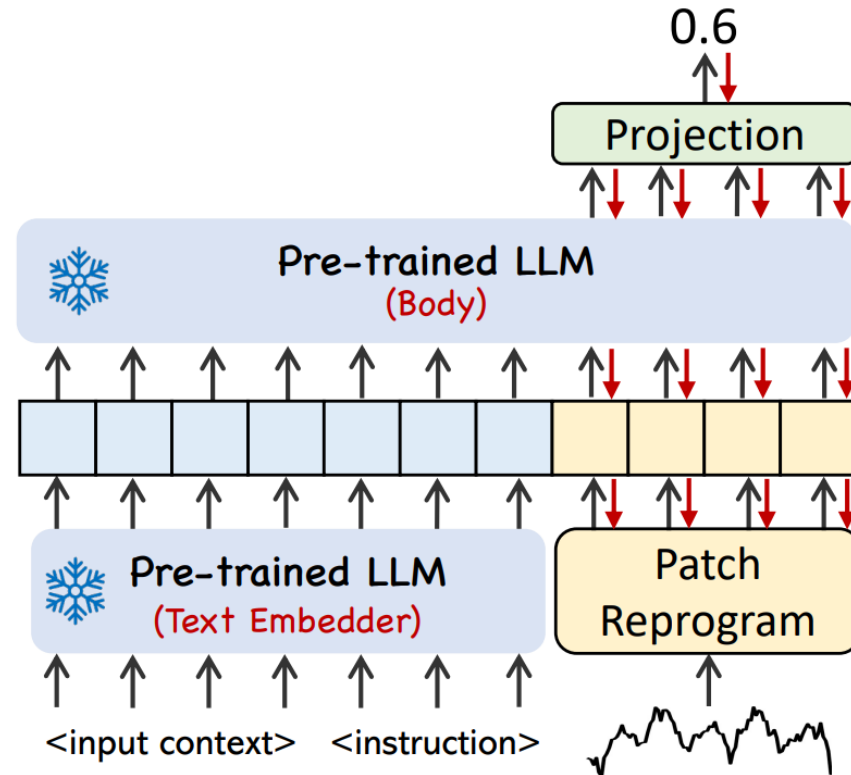
앞서 만들었던 두 토큰을
concat 하여 backbone에
입력

TIME-LLM

❖ Output Projection



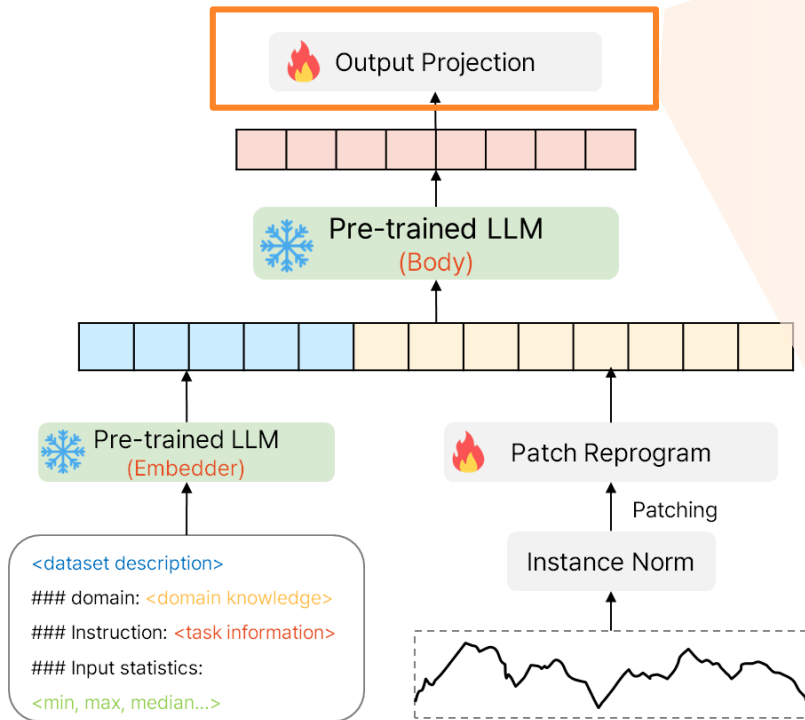
Output Projection



Prefix 부분을 제거한
시계열 토큰 부분만
projection

TIME-LLM

❖ Output Projection



Forecasts



Flatten & Linear

패치들을 flatten 한 후
예측

Output Patch
Embeddings

TIME-LLM

❖ Experiments

- Backbone: Llama-7B
- Metric: MSE, MAE (Long) / SMAPE, MSAE, OWA (Short)

Tasks	Dataset	Dim.	Series Length	Dataset Size	Frequency	Domain
Long-term Forecasting	ETTm1	7	{96, 192, 336, 720}	(34465, 11521, 11521)	15 min	Temperature
	ETTm2	7	{96, 192, 336, 720}	(34465, 11521, 11521)	15 min	Temperature
	ETTh1	7	{96, 192, 336, 720}	(8545, 2881, 2881)	1 hour	Temperature
	ETTh2	7	{96, 192, 336, 720}	(8545, 2881, 2881)	1 hour	Temperature
	Electricity	321	{96, 192, 336, 720}	(18317, 2633, 5261)	1 hour	Electricity
	Traffic	862	{96, 192, 336, 720}	(12185, 1757, 3509)	1 hour	Transportation
	Weather	21	{96, 192, 336, 720}	(36792, 5271, 10540)	10 min	Weather
	ILI	7	{24, 36, 48, 60}	(617, 74, 170)	1 week	Illness
Short-term Forecasting	M3-Quarterly	1	8	(756, 0, 756)	Quarterly	Multiple
	M4-Yearly	1	6	(23000, 0, 23000)	Yearly	Demographic
	M4-Quarterly	1	8	(24000, 0, 24000)	Quarterly	Finance
	M4-Monthly	1	18	(48000, 0, 48000)	Monthly	Industry
	M4-Weakly	1	13	(359, 0, 359)	Weakly	Macro
	M4-Daily	1	14	(4227, 0, 4227)	Daily	Micro
	M4-Hourly	1	48	(414, 0, 414)	Hourly	Other

TIME-LLM

❖ Experiments

- Predict length{96, 192, 336, 720} 평균값
- 기존 SOTA 모델인 PatchTST 보다 낮은 오차를 보이며, LLM을 사용하는 GPT4TS보다 성능 향상을 보임

Long-term
Forecasting

Methods	TIME-LLM (Ours)		GPT4TS (2023a)		DLinear (2023)		PatchTST (2023)		TimesNet (2023)		FEDformer (2022)		Autoformer (2021)		Stationary (2022)		ETSformer (2022)		LightTS (2022a)		Informer (2021)		Reformer (2020)	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
<i>ETTh1</i>	0.408	0.423	0.465	0.455	0.422	0.437	<u>0.413</u>	<u>0.430</u>	0.458	0.450	0.440	0.460	0.496	0.487	0.570	0.537	0.542	0.510	0.491	0.479	1.040	0.795	1.029	0.805
<i>ETTh2</i>	<u>0.334</u>	<u>0.383</u>	0.381	0.412	0.431	0.446	0.330	0.379	0.414	0.427	0.437	0.449	0.450	0.459	0.526	0.516	0.439	0.452	0.602	0.543	4.431	1.729	6.736	2.191
<i>ETTm1</i>	0.329	0.372	0.388	0.403	0.357	<u>0.378</u>	<u>0.351</u>	0.380	0.400	0.406	0.448	0.452	0.588	0.517	0.481	0.456	0.429	0.425	0.435	0.437	0.961	0.734	0.799	0.671
<i>ETTm2</i>	0.251	0.313	0.284	0.339	0.267	0.333	<u>0.255</u>	<u>0.315</u>	0.291	0.333	0.305	0.349	0.327	0.371	0.306	0.347	0.293	0.342	0.409	0.436	1.410	0.810	1.479	0.915
<i>Weather</i>	0.225	0.257	0.237	0.270	0.248	0.300	0.225	<u>0.264</u>	0.259	0.287	0.309	0.360	0.338	0.382	0.288	0.314	0.271	0.334	0.261	0.312	0.634	0.548	0.803	0.656
<i>ECL</i>	0.158	0.252	0.167	0.263	0.166	0.263	<u>0.161</u>	0.252	0.192	0.295	0.214	0.327	0.227	0.338	0.193	0.296	0.208	0.323	0.229	0.329	0.311	0.397	0.338	0.422
<i>Traffic</i>	0.388	<u>0.264</u>	0.414	0.294	0.433	0.295	<u>0.390</u>	0.263	0.620	0.336	0.610	0.376	0.628	0.379	0.624	0.340	0.621	0.396	0.622	0.392	0.764	0.416	0.741	0.422
<i>ILI</i>	1.435	<u>0.801</u>	1.925	0.903	2.169	1.041	<u>1.443</u>	0.797	2.139	0.931	2.847	1.144	3.006	1.161	2.077	0.914	2.497	1.004	7.382	2.003	5.137	1.544	4.724	1.445
1 st Count	7		0		0		<u>5</u>		0		0		0		0		0		0		0		0	

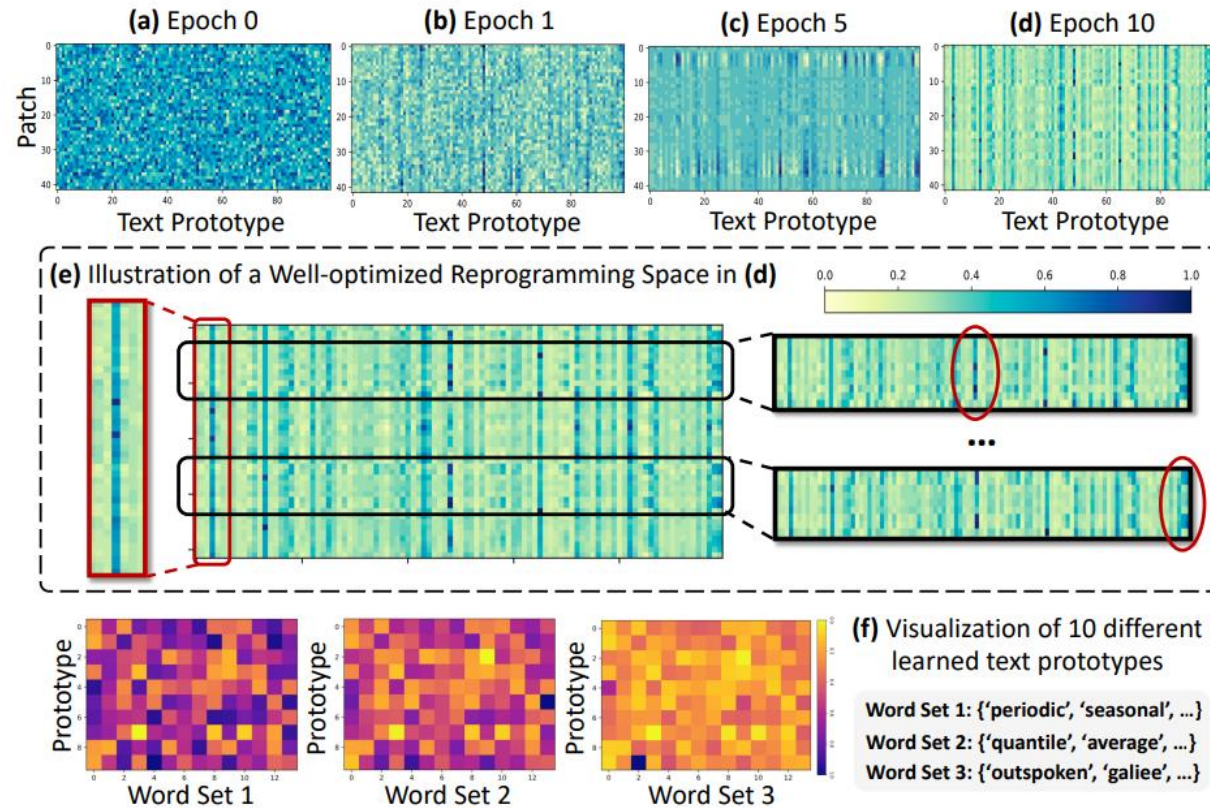
Short-term
Forecasting

Methods	TIME-LLM (Ours)		GPT4TS (2023a)	TimesNet (2023)	PatchTST (2023)	N-HiTS (2023b)	N-BEATS (2020)	ETSformer (2022)	LightTS (2022a)	DLinear (2023)	FEDformer (2022)	Stationary (2022)	Autoformer (2021)	Informer (2021)	Reformer (2020)
	SMAPE	MASE													
Average SMAPE	11.983		12.69	12.88	12.059	<u>12.035</u>	12.25	14.718	13.525	13.639	13.16	12.780	12.909	14.086	18.200
Average MASE	1.595		1.808	1.836	1.623	<u>1.625</u>	1.698	2.408	2.111	2.095	1.775	1.756	1.771	2.718	4.223
Average OWA	0.859		0.94	0.955	<u>0.869</u>	<u>0.869</u>	0.896	1.172	1.051	1.051	0.949	0.930	0.939	1.230	1.775

TIME-LLM

❖ Reprogramming Interpretation

- 학습이 진행될수록 특정 text prototypes만 사용되고, 각 패치들은 그 안에서 다양한 text를 선택하며 mapping 되고 있음
- Mapping 되는 단어들이 실제 시계열 패치를 잘 설명하고 있는지는 검증이 필요해 보임



TIME-FFM

❖ TIME-FFM: Towards LM-Empowered Federated Foundation Model for Time Series Forecasting (2024, NeurIPS)

- 사전 학습된 LM(Language Model)을 활용하여 시계열 예측을 위한 Federated Foundation Model 제안
- 중앙 집중형 모델 구축 과정에서 데이터 프라이버시 문제를 해결하기 위해 연합학습 도입

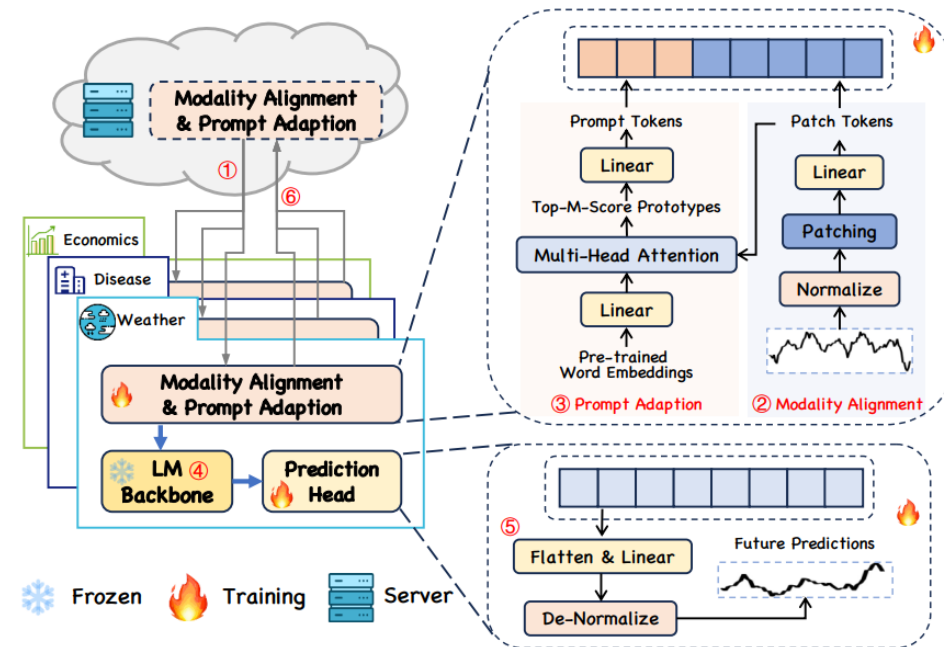
TIME-FFM: Towards LM-Empowered Federated Foundation Model for Time Series Forecasting

Qingxiang Liu^{1,2} Xu Liu³ Chenghao Liu⁴ Qingsong Wen⁵ Yuxuan Liang^{1*}

¹ The Hong Kong University of Science and Technology (Guangzhou)
² Institute of Computing Technology Chinese Academy of Sciences
³ National University of Singapore ⁴ Salesforce AI Research ⁵ Squirrel AI
qingxiangliu737@gmail.com, liuxu@comp.nus.edu.sg
chenghao.liu@salesforce.com, qingsongedu@gmail.com, yuxliang@outlook.com

Abstract

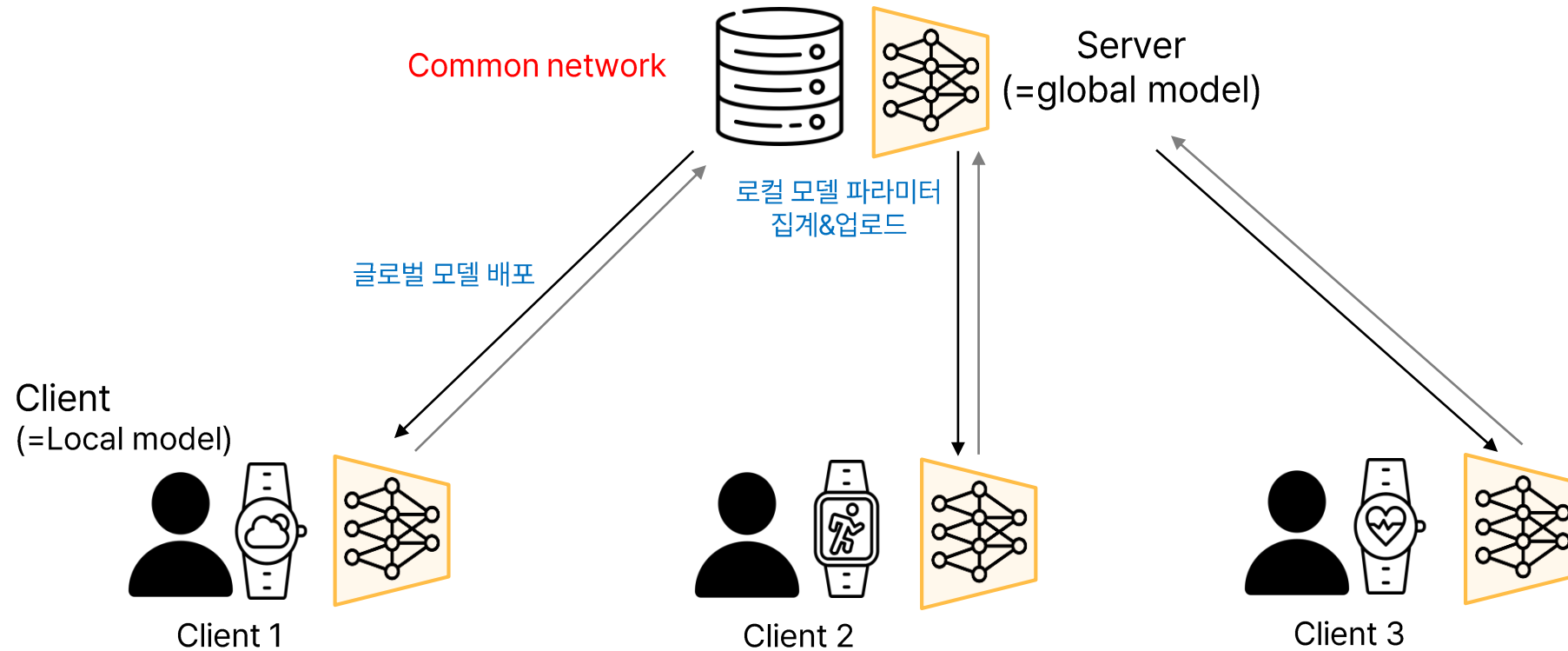
Unlike natural language processing and computer vision, the development of Foundation Models (FMs) for time series forecasting is blocked due to data scarcity. While recent efforts are focused on building such FMs by unlocking the potential of language models (LMs) for time series analysis, dedicated parameters for various downstream forecasting tasks need training, which hinders the common knowledge sharing across domains. Moreover, data owners may hesitate to share the access to local data due to privacy concerns and copyright protection, which makes it impossible to simply construct a FM on cross-domain training instances. To address these issues, we propose TIME-FFM, a Federated Foundation Model for TIME series forecasting by leveraging pretrained LMs. Specifically, we begin by transforming time series into the modality of text tokens. To bootstrap LMs for time series reasoning, we propose a prompt adaption module to determine domain-customized prompts dynamically instead of artificially. Given the data heterogeneity across domains, we design a personalized federated training strategy by learning global encoders and local prediction heads. Our comprehensive experiments indicate that TIME-FFM outperforms state-of-the-arts and promises effective few-shot and zero-shot forecaster. The code is available at <https://github.com/CityMind-Lab/NeurIPS24-Time-FFM/tree/main>



Background

❖ 연합학습이란?

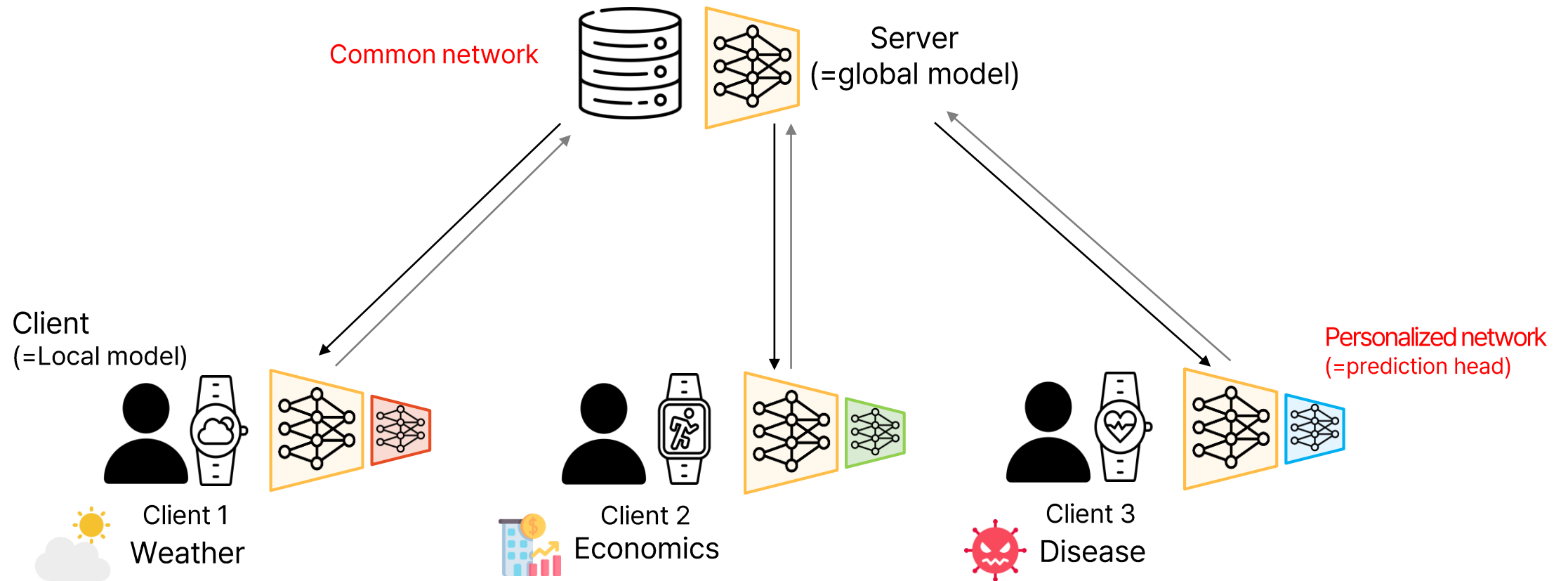
- 데이터를 직접 공유하지 않고, 로컬에서 학습한 **파라미터만** 집계하여 글로벌 모델 구축 → **프라이버시 보호**



Background

❖ 개인화 연합학습이란?

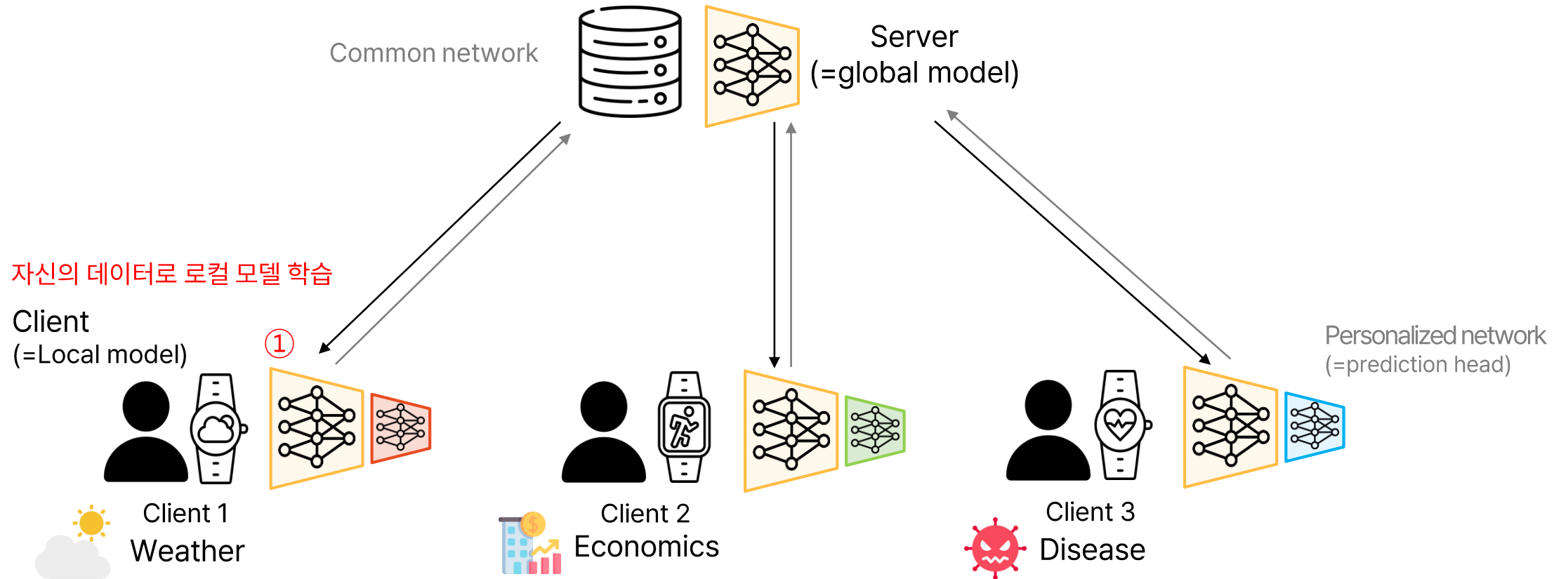
- 클라이언트 간 이질성을 고려하여, 공통 지식을 공유하면서 각 클라이언트에 맞춤형 모델을 학습 → 프라이버시 보호 + 개별 클라이언트 특성 반영



Background

❖ 개인화 연합학습이란?

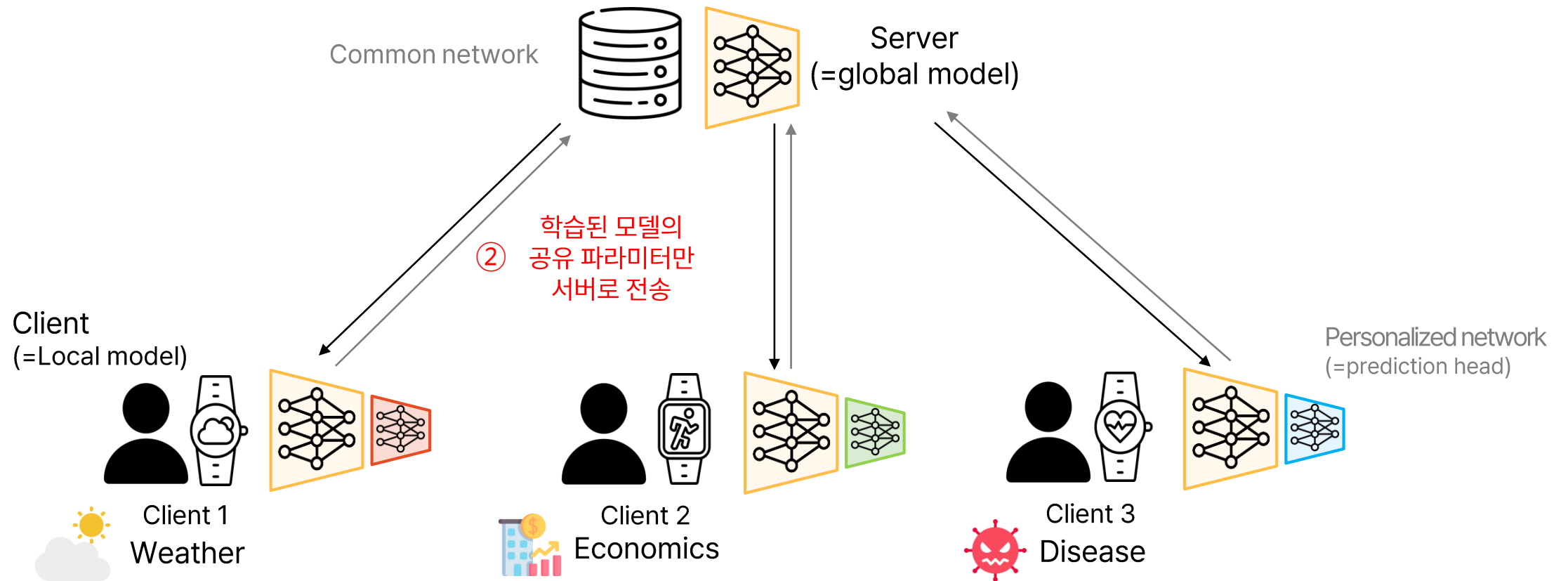
- 클라이언트 간 이질성을 고려하여, 공통 지식을 공유하면서 각 클라이언트에 맞춤형 모델을 학습 → 프라이버시 보호 + 개별 클라이언트 특성 반영



Background

❖ 개인화 연합학습이란?

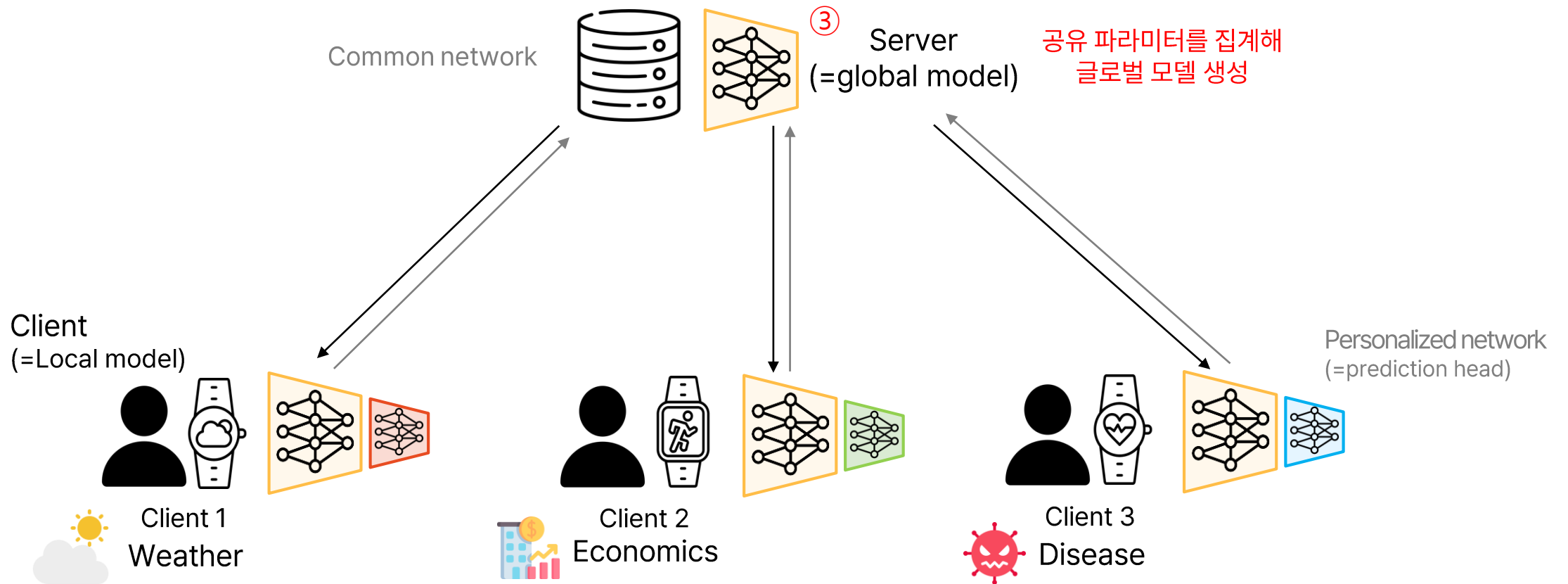
- 클라이언트 간 이질성을 고려하여, 공통 지식을 공유하면서 각 클라이언트에 맞춤형 모델을 학습 → 프라이버시 보호 + 개별 클라이언트 특성 반영



Background

❖ 개인화 연합학습이란?

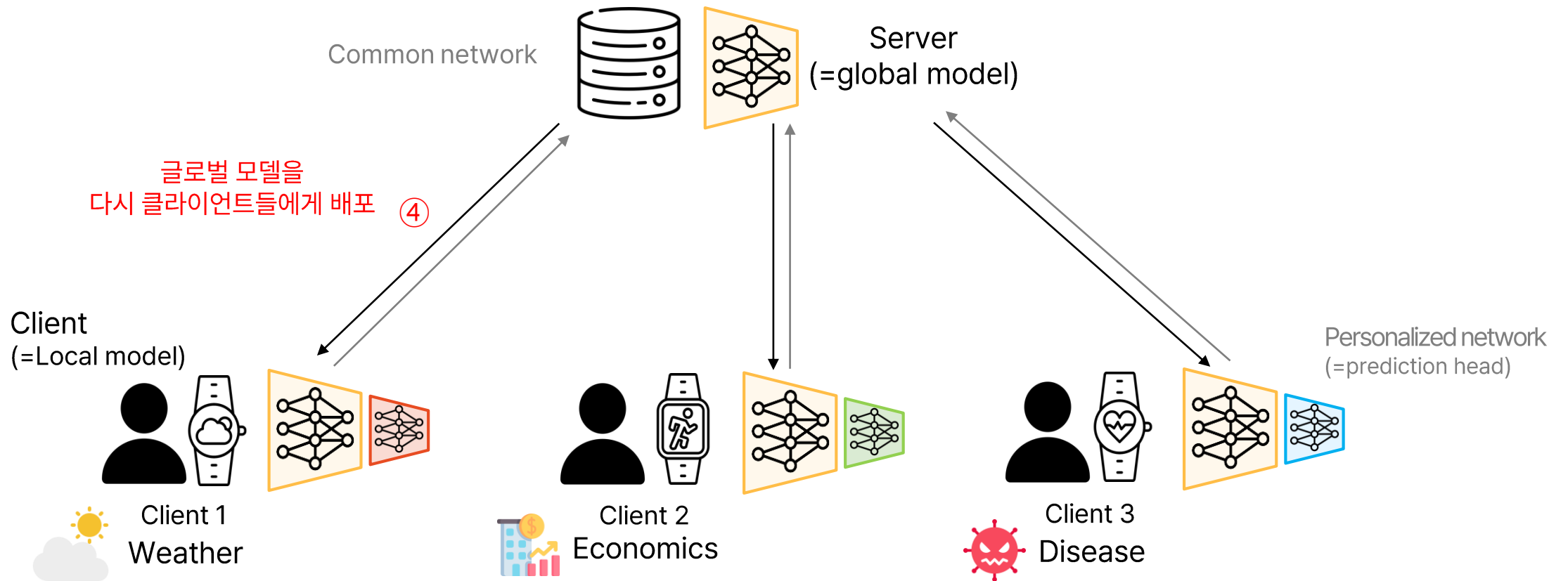
- 클라이언트 간 이질성을 고려하여, 공통 지식을 공유하면서 각 클라이언트에 맞춤형 모델을 학습 → 프라이버시 보호 + 개별 클라이언트 특성 반영



Background

❖ 개인화 연합학습이란?

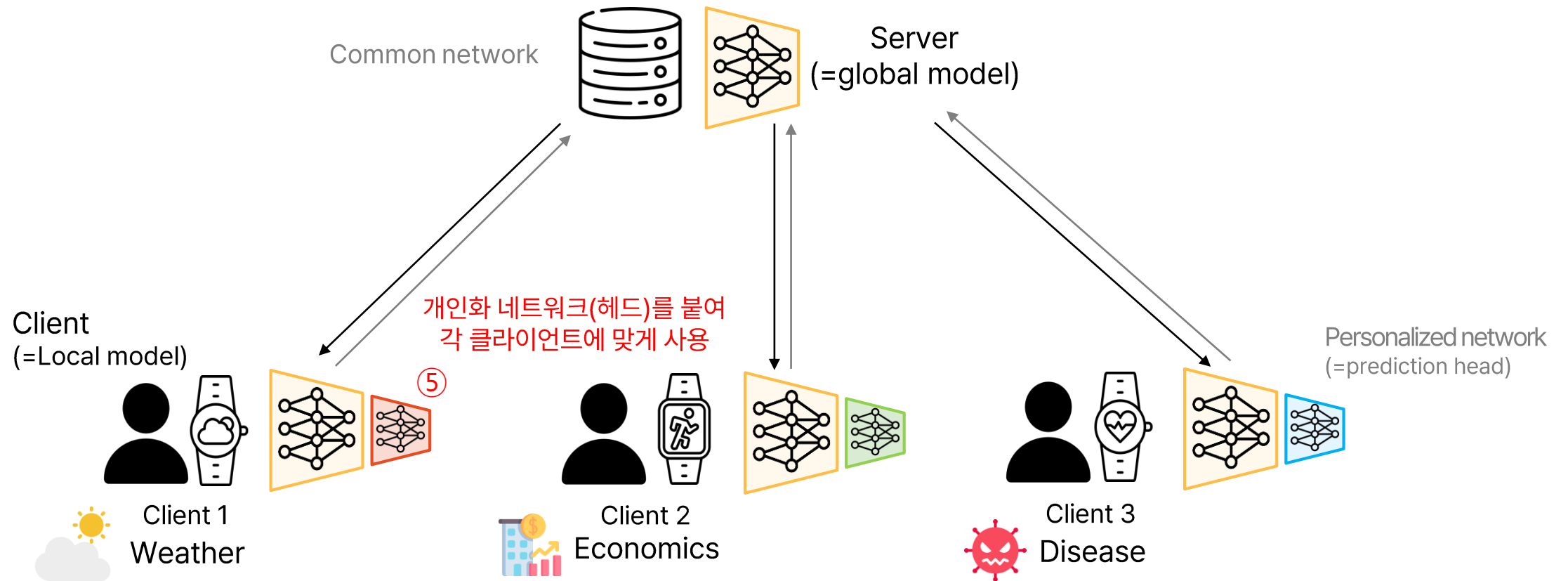
- 클라이언트 간 이질성을 고려하여, 공통 지식을 공유하면서 각 클라이언트에 맞춤형 모델을 학습 → 프라이버시 보호 + 개별 클라이언트 특성 반영



Background

❖ 개인화 연합학습이란?

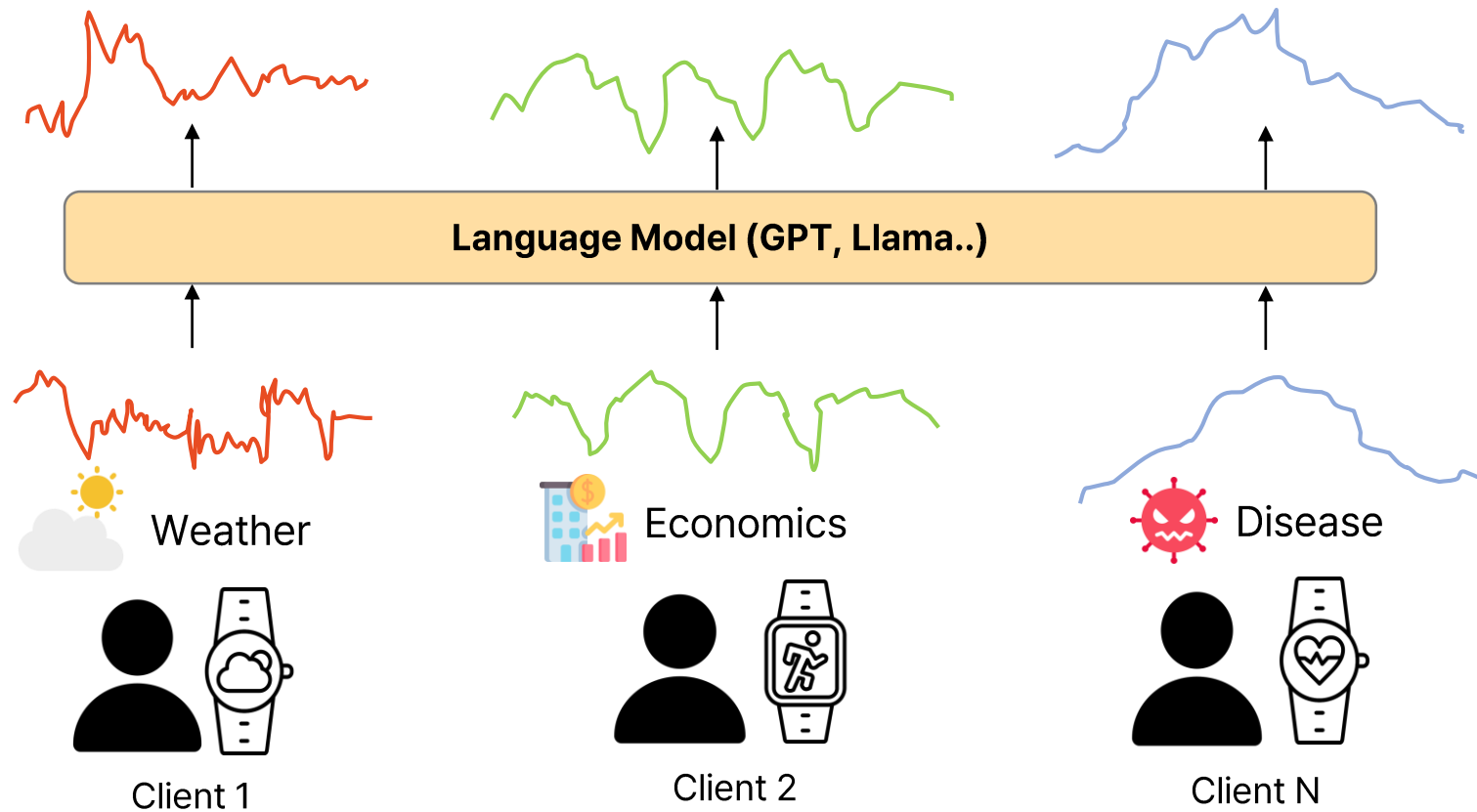
- 클라이언트 간 이질성을 고려하여, 공통 지식을 공유하면서 각 클라이언트에 맞춤형 모델을 학습 → 프라이버시 보호 + 개별 클라이언트 특성 반영



Background

❖ 시계열 예측을 위한 연합학습

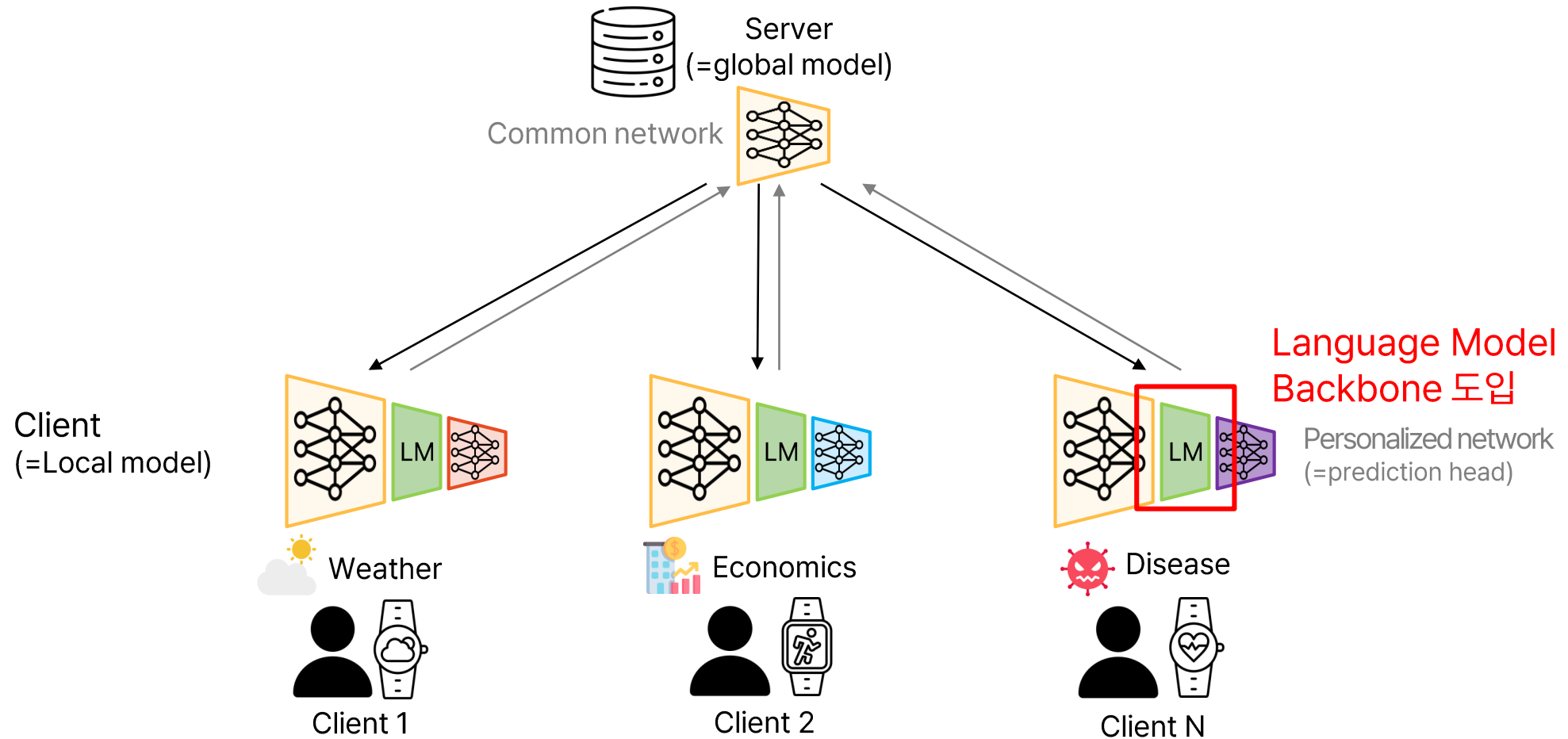
- 성능이 좋은 사전 학습된 LM(Language Model)을 활용한 시계열 예측 Foundation Model 연구 활발



TIME-FFM

❖ TIME-FFM: Towards LM-Empowered Federated Foundation Model for Time Series Forecasting (2024, NeurIPS)

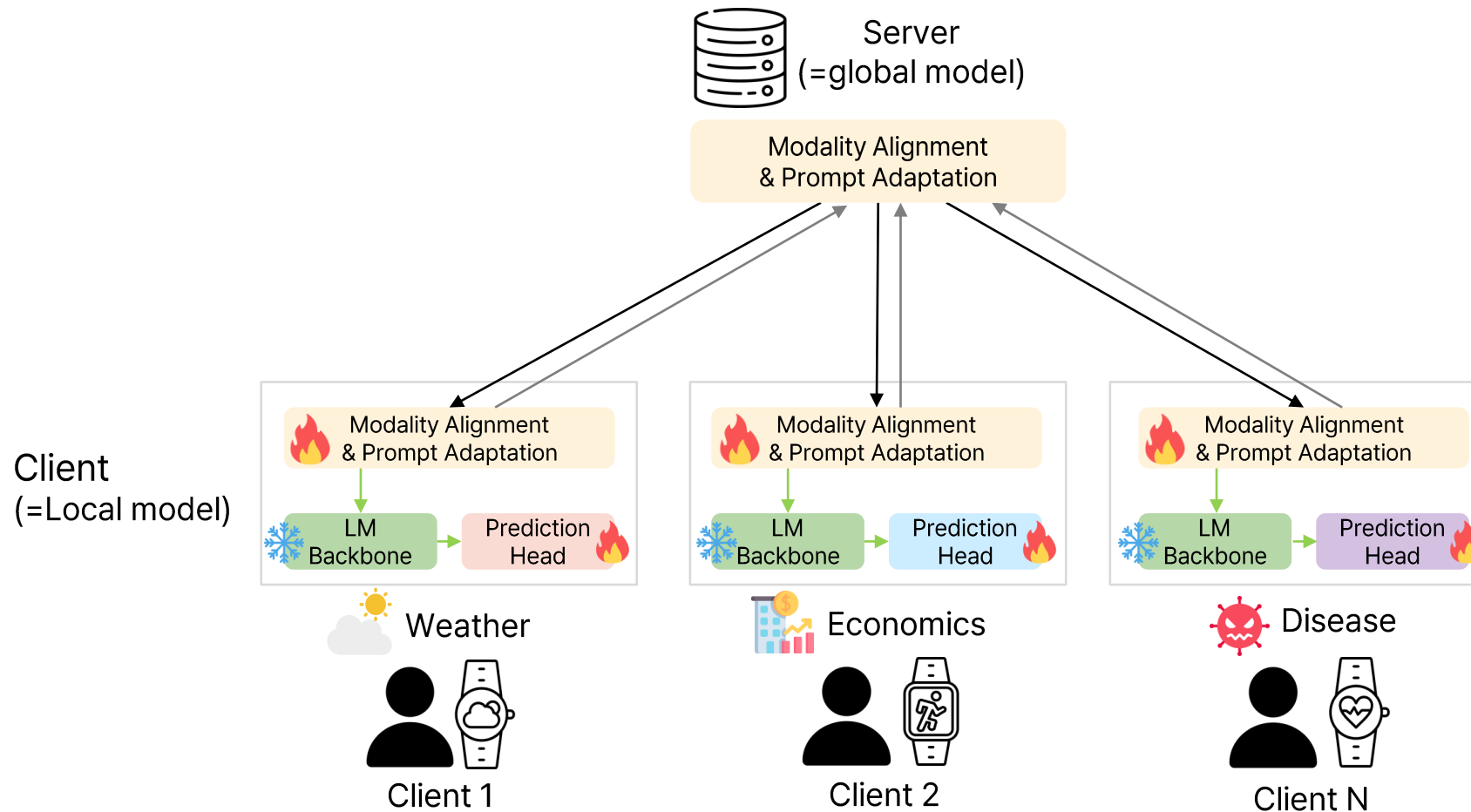
- 사전 학습된 LM(Language Model)을 활용하여 시계열 예측을 위한 Federated Foundation Model 제안



TIME-FFM

❖ TIME-FFM: Towards LM-Empowered Federated Foundation Model for Time Series Forecasting (2024, NeurIPS)

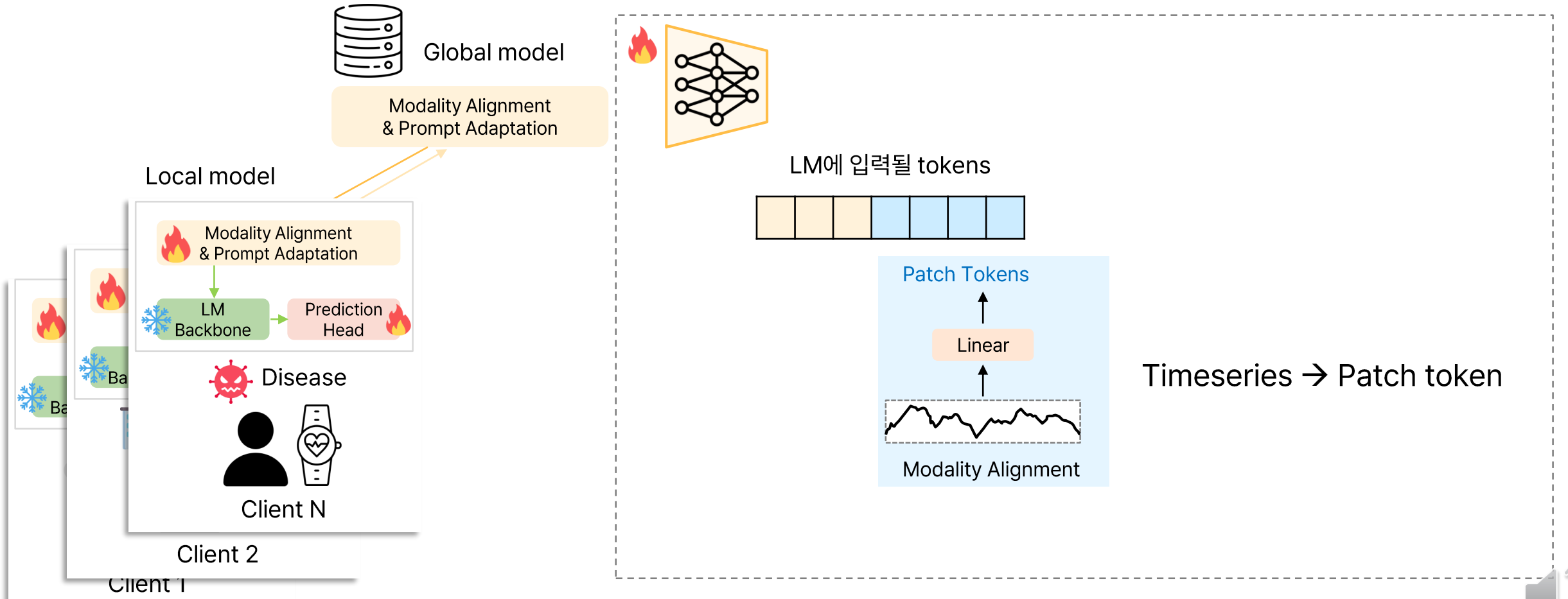
- 사전 학습된 LM(Language Model)을 활용하여 시계열 예측을 위한 Federated Foundation Model 제안



TIME-FFM

❖ Alignment Modality

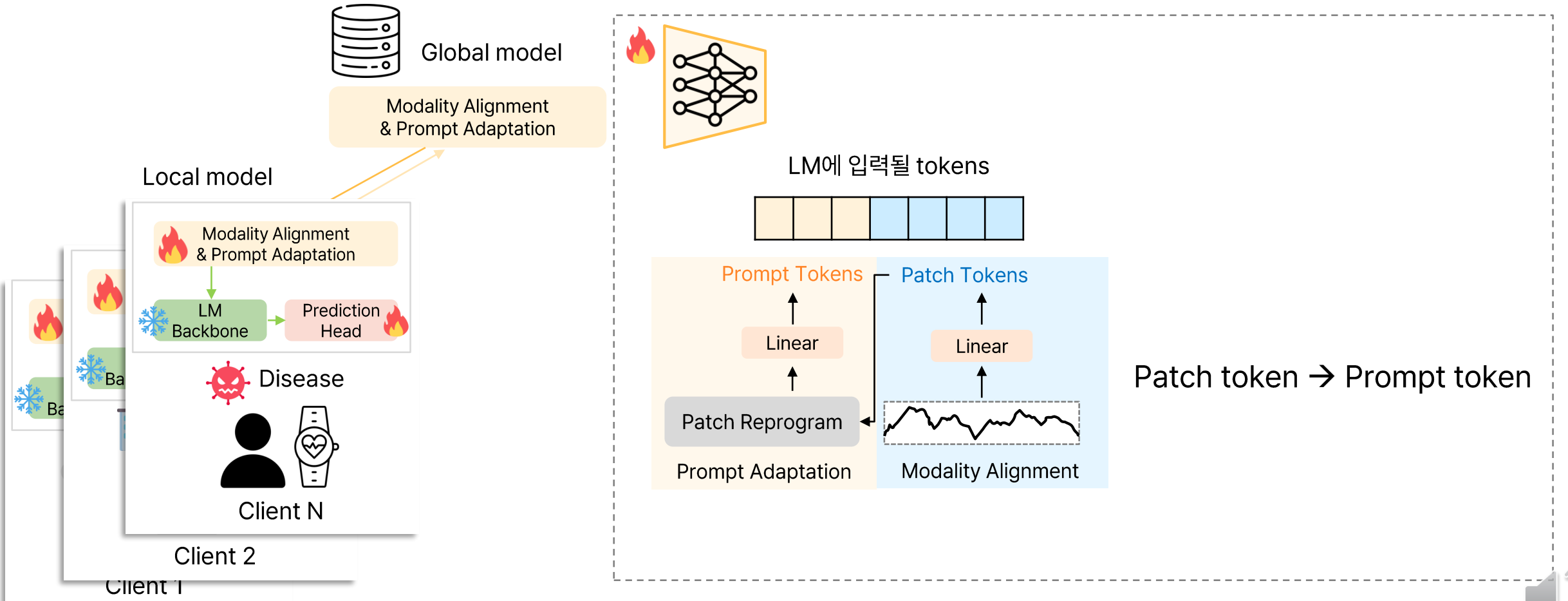
- 시계열을 패치 단위로 나눠 선형 레이어를 통해 LM 백본의 입력 차원에 맞는 토큰으로 투영



TIME-FFM

❖ Prompt Adaptation

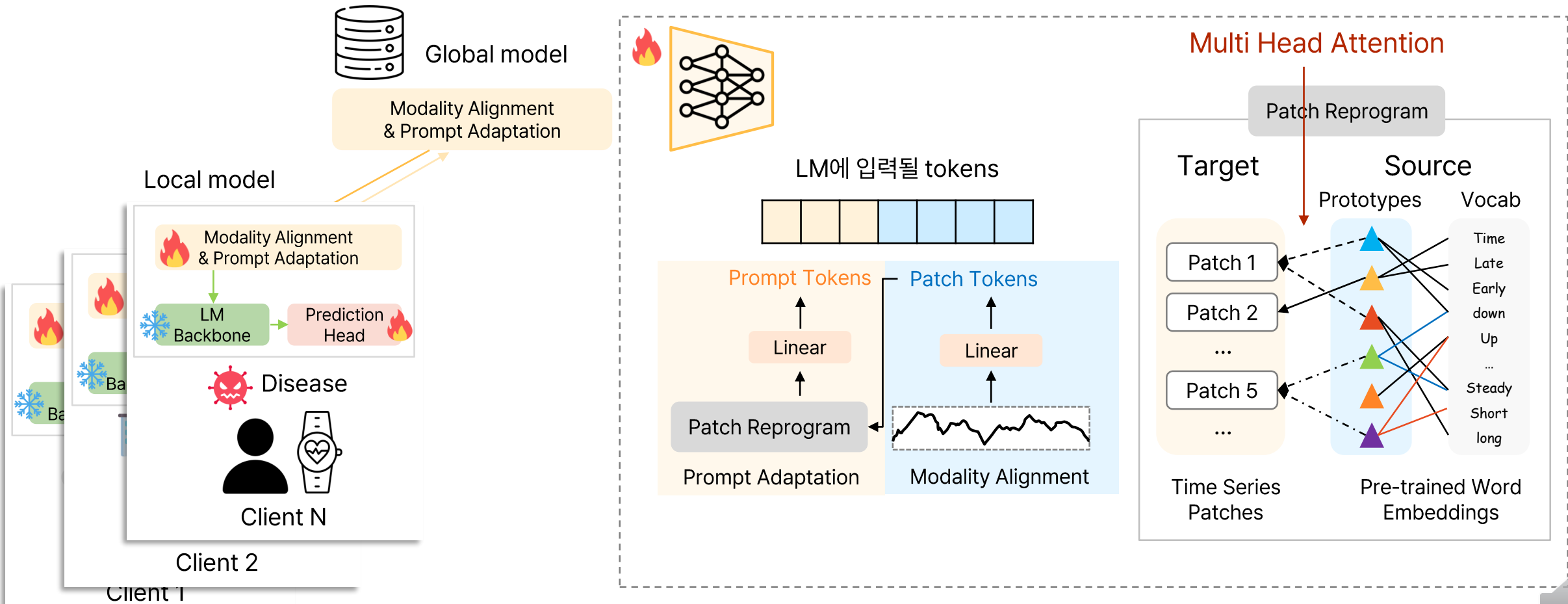
- 패치 토큰과 사전 학습된 LM 기반으로 **도메인별 프롬프트**를 동적으로 결정



TIME-FFM

❖ Prompt Adaptation

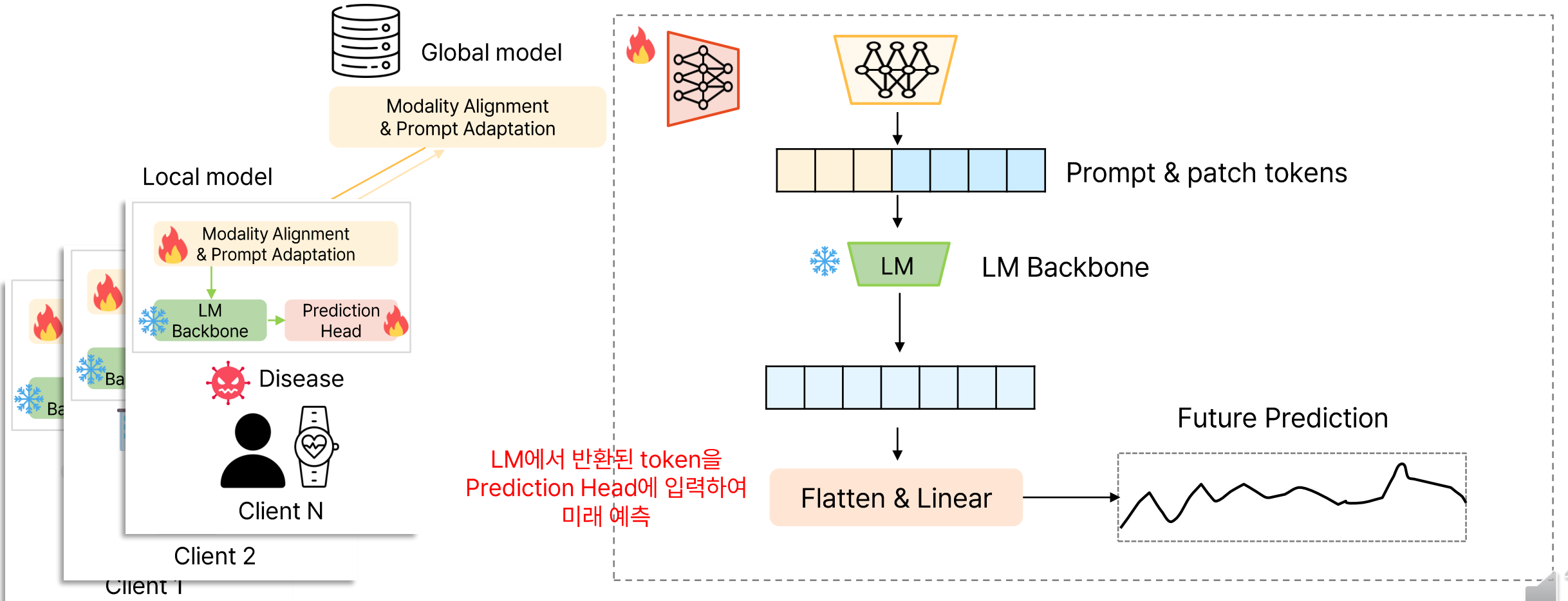
- 패치 토큰과 사전 학습된 LM 기반으로 **도메인별 프롬프트**를 동적으로 결정



TIME-FFM

❖ Forecasting

- 개인화 네트워크인 예측 헤드를 통해 각 클라이언트별 시계열 예측



LM에서 반환된 token을
Prediction Head에 입력하여
미래 예측

TIME-FFM

❖ Experiments

- 8개의 데이터셋을 다른 도메인의 클라이언트로 가정
- Backbone: GPT-2
- Metric: MSE, MAE

Dataset	c_i	Dataset Size	Batch Size	OverSampling Times	Frequency	Application Domain
ETTh1	7	(8545, 2881, 2881)	32	-	1 hour	Electrical Asset Monitoring
ETTh2	7	(8545, 2881, 2881)	32	-	1 hour	Electrical Asset Monitoring
ETTh1	7	(34465, 11521, 11521)	64	-	15 minutes	Electrical Asset Monitoring
ETTh2	7	(34465, 11521, 11521)	64	-	15 minutes	Electrical Asset Monitoring
Electricity	321	(18317, 2633, 5261)	24	-	1 hour	Energy Consumption
Weather	21	(36792, 5271, 10540)	64	-	10 minutes	Weather Forecasting
Exchange	8	(5120, 665, 1422)	24	-	1 day	International Trade
ILI	7	(617, 74, 170)	16	12	1 week	Illness Monitoring

TIME-FFM

❖ Experiments

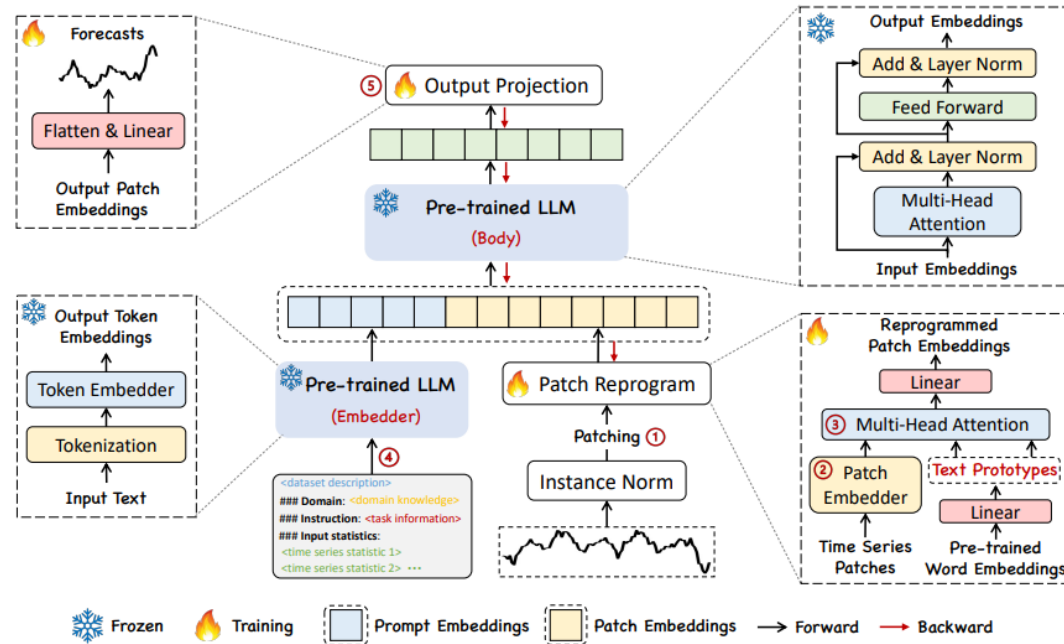
- Predict length{96, 192, 336, 720} 평균값
- TY1 (federated fine-tuning methods); TY2 (across-dataset centralized methods); TY3 (dataset-specific centralized methods)
- 연합학습 상황임에도 불구하고 중앙 집중형 모델(UniTime)보다 더 낮거나 준수한 성능을 보임

Type	TY1								TY2				TY3										
Method	TIME-FFM	FedIT	FedAdapter ^H	FedAdapter ^P	UniTime	GPT4TS	PatchTST	TimesNet	DLinear	FEDformer	Autoformer	Informer											
Metric	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE	MSE MAE
ETTh1	0.442 0.434	0.481 0.461	0.488 0.467	0.503 0.479	0.442 0.448	0.502 0.461	0.472 0.451	0.458 0.450	0.456 0.452	0.440 0.460	0.496 0.487	1.040 0.795											
ETTh2	0.382 0.406	0.374 0.396	0.373 0.398	0.380 0.403	0.378 0.403	0.386 0.406	0.398 0.416	0.414 0.427	0.559 0.515	0.437 0.449	0.450 0.459	4.431 1.729											
ETTm1	0.399 0.402	0.644 0.517	0.643 0.511	0.640 0.516	0.385 0.399	0.551 0.483	0.971 0.629	0.383 0.406	0.403 0.407	0.448 0.452	0.588 0.517	0.961 0.734											
ETTm2	0.286 0.332	0.297 0.341	0.295 0.340	0.298 0.342	0.293 0.334	0.321 0.356	0.340 0.373	0.291 0.322	0.350 0.401	0.305 0.349	0.327 0.371	1.410 0.810											
Electricity	0.216 0.299	0.390 0.478	0.408 0.489	0.334 0.420	0.216 0.305	0.251 0.338	0.221 0.311	0.193 0.295	0.212 0.300	0.214 0.327	0.227 0.338	0.311 0.397											
Weather	0.270 0.288	0.282 0.310	0.282 0.308	0.287 0.309	0.253 0.276	0.293 0.309	0.304 0.323	0.259 0.287	0.265 0.317	0.309 0.360	0.338 0.382	0.634 0.548											
Exchange	0.338 0.391	0.389 0.423	0.382 0.419	0.380 0.417	0.364 0.404	0.421 0.446	0.411 0.444	0.416 0.443	0.354 0.414	0.519 0.500	0.613 0.539	1.550 0.998											
ILI	2.107 0.924	4.423 1.448	5.247 1.621	5.251 1.600	2.137 0.929	3.678 1.372	4.210 1.480	2.139 0.931	2.616 1.090	2.847 1.144	3.006 1.161	5.137 1.544											
Average	0.555 0.434	0.910 0.547	1.015 0.569	1.009 0.561	0.559 0.437	0.800 0.521	0.916 0.553	0.569 0.445	0.652 0.487	0.690 0.505	0.756 0.532	1.934 0.944											
1 st Count	8	1	1	0	3	0	0	4	0	1	0	0											

Conclusion

❖ Conclusion

- 사전 학습된 LLM을 기반으로 **시계열과 언어의 모달리티를 정렬**하는 방법을 제안 (TIME-LLM)
- 최근 **데이터 프라이버시** 문제가 대두되면서, 데이터가 아닌 파라미터만 공유하여 모델을 구축하는 **연합학습** 상황으로 확장 (TIME-FFM)
- LLM을 backbone으로 사용하는 것에서 더 나아가, **사전 학습된 단어**들을 시계열과 연관 지어 활용하는 것이 인상적
- 하지만, 시계열에 대해 충분히 **설명력** 있는 단어가 연결되고 있는지에 대한 **분석**이 더 필요해 보임



Thank You

