

# 네이버 수요예측의 비밀 HyperCLOVA, 빅데이터에서 예측 시그널을 찾을 수 있는 이유

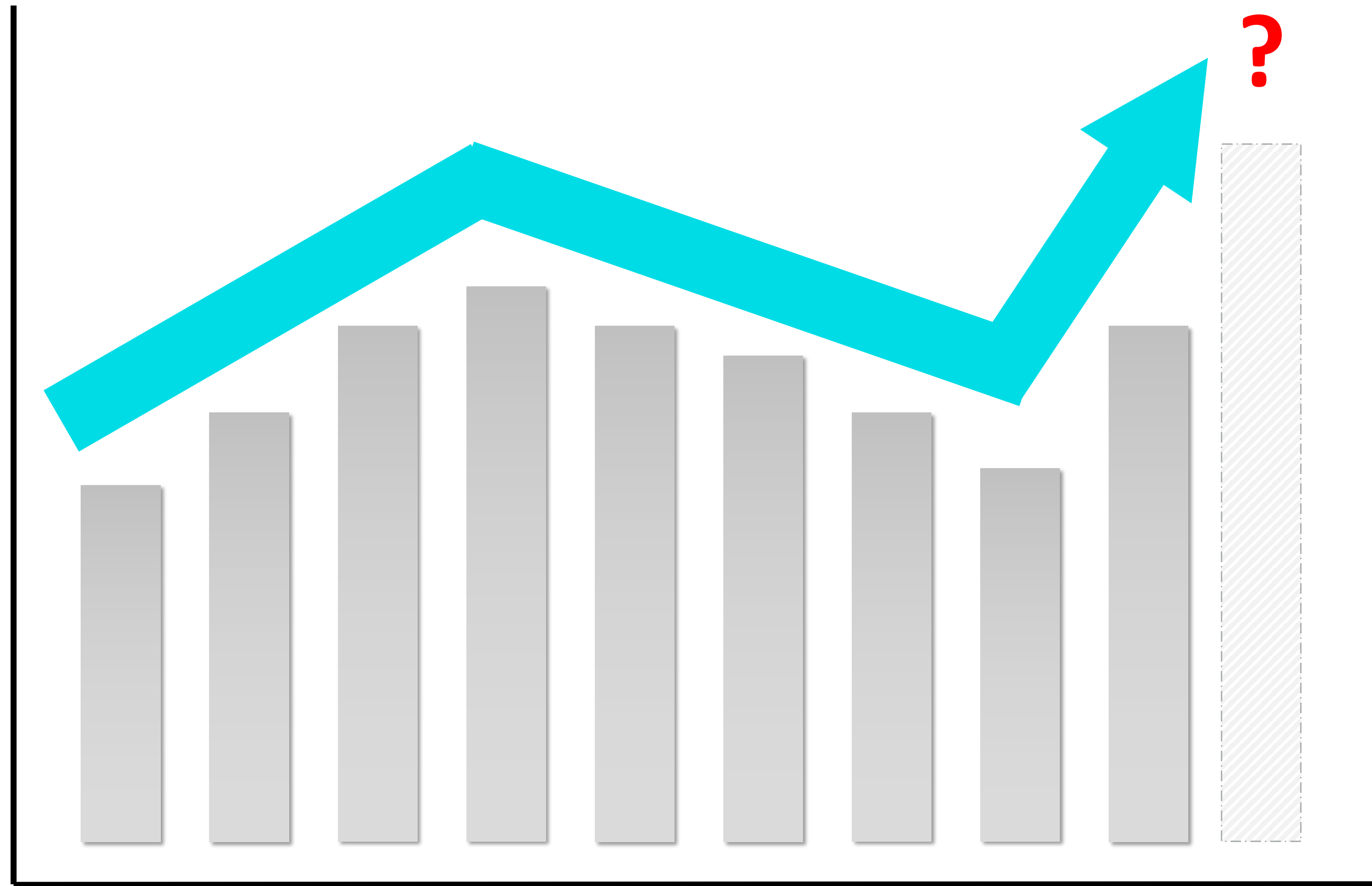
김경민, 박영진, 이주호, 홍영택 NAVER CLOVA

# CONTENTS

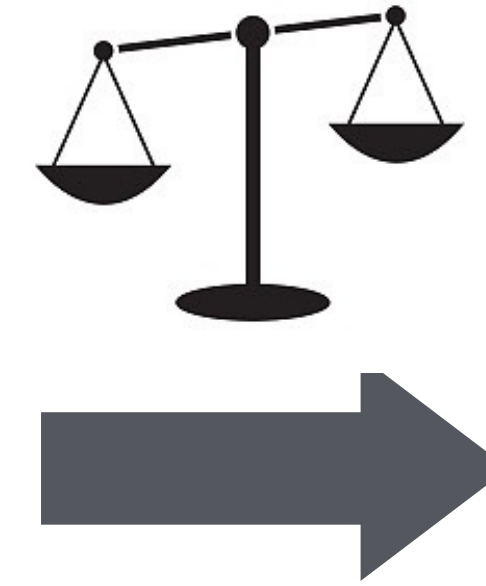
1. 수요예측용 HyperCLOVA는 무엇인가?
2. 기존 방식을 사용한 수요예측
3. HyperCLOVA 수요예측
4. HyperCLOVA 활용 결과 (SKU단위 수요예측)

# 1. 수요예측용 HyperCLOVA는 무엇인가?

# 1.1 상품 수요예측과 의사결정



수요예측



**물류회사**  
“내일 수요가 xx니까  
패킹 인력을 xx명 고용해야지”

**판매업자**  
“일주일동안 수요가 xx니까  
물건을 미리 떼와야겠구나”

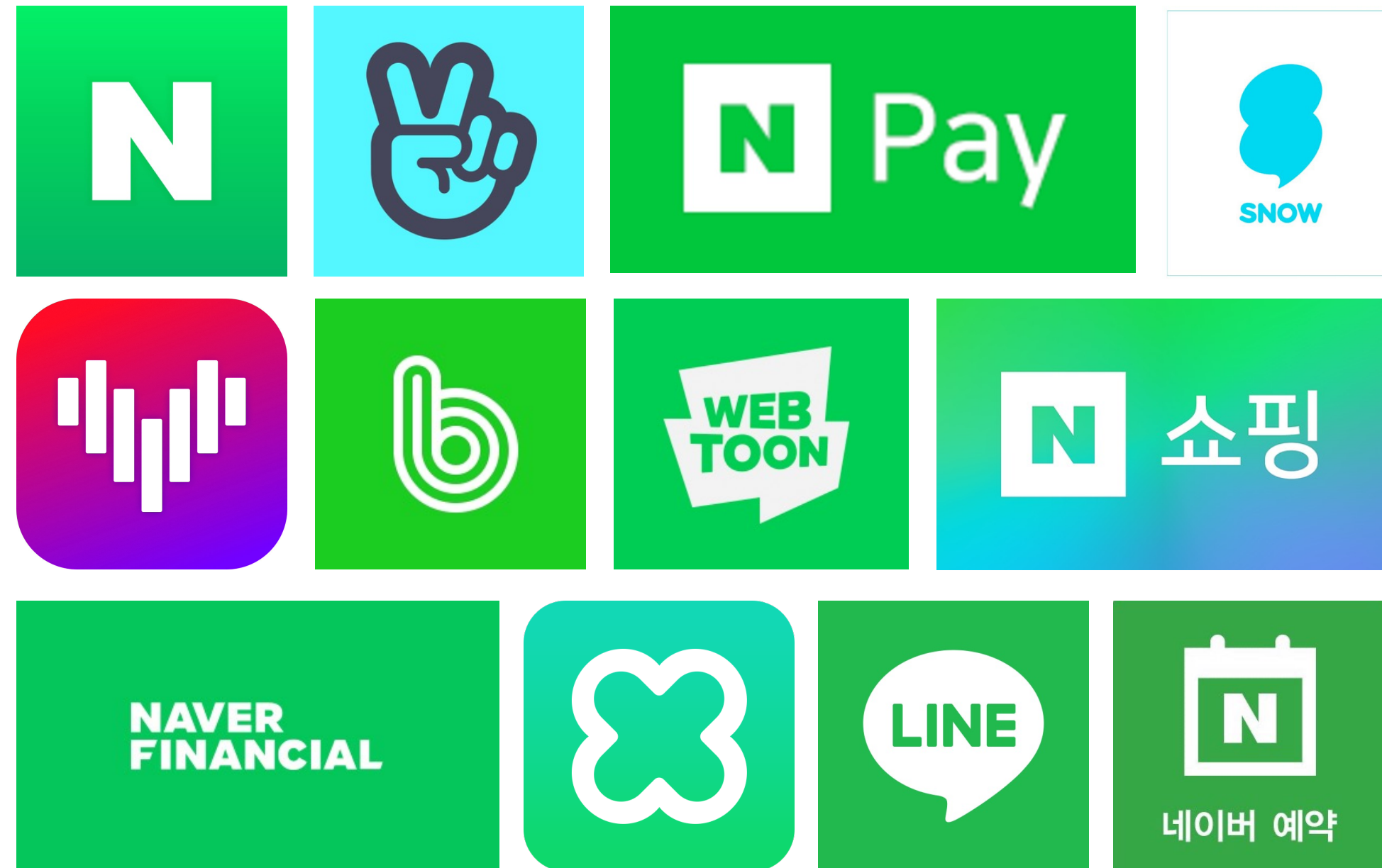
예측기반 의사결정

# 1.2 네이버 상품 수요예측의 장점

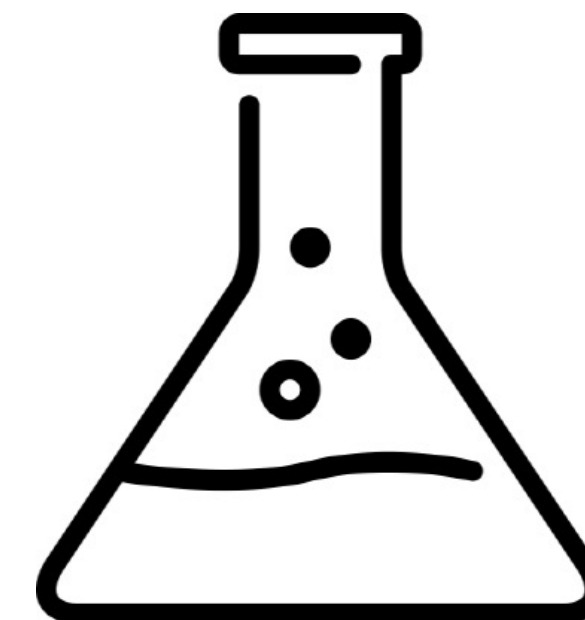
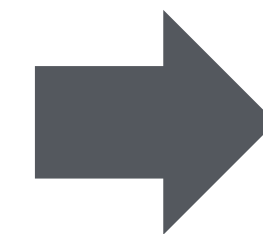


# 1.3 큰 규모의 네이버 데이터 활용은 쉽지 않다

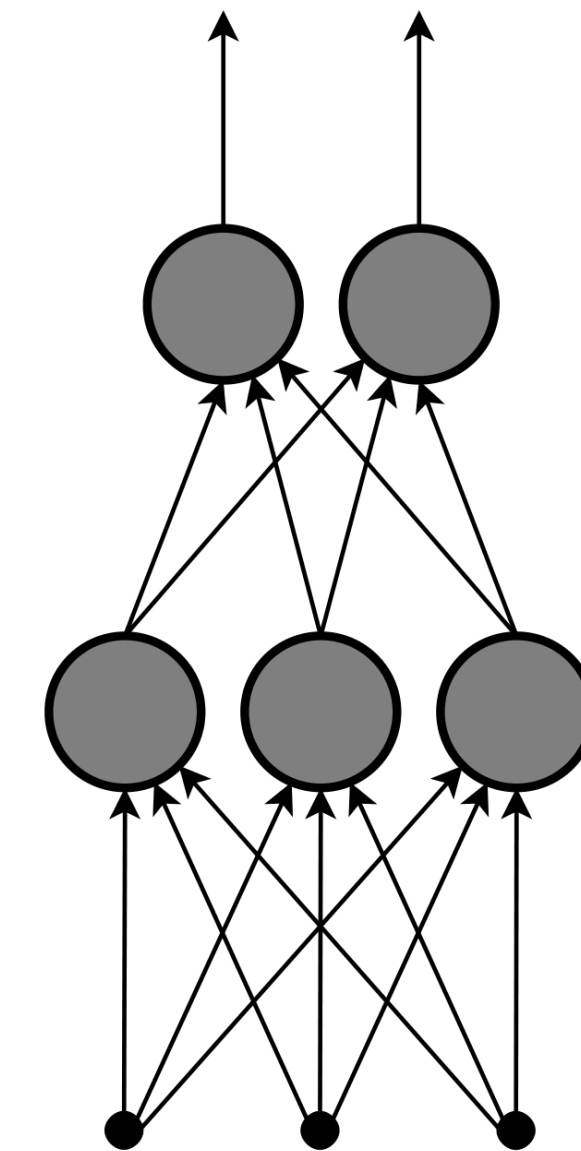
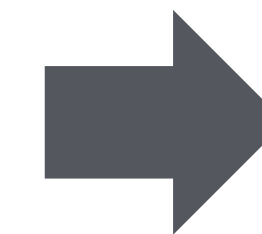
#대규모\_데이터\_분석 #낮은\_SNR #노가다 #시간효율성



빅데이터



피처검증



모델 사용

# 1.4 상황에 딱 맞는 모델을 자동으로 만들어주는

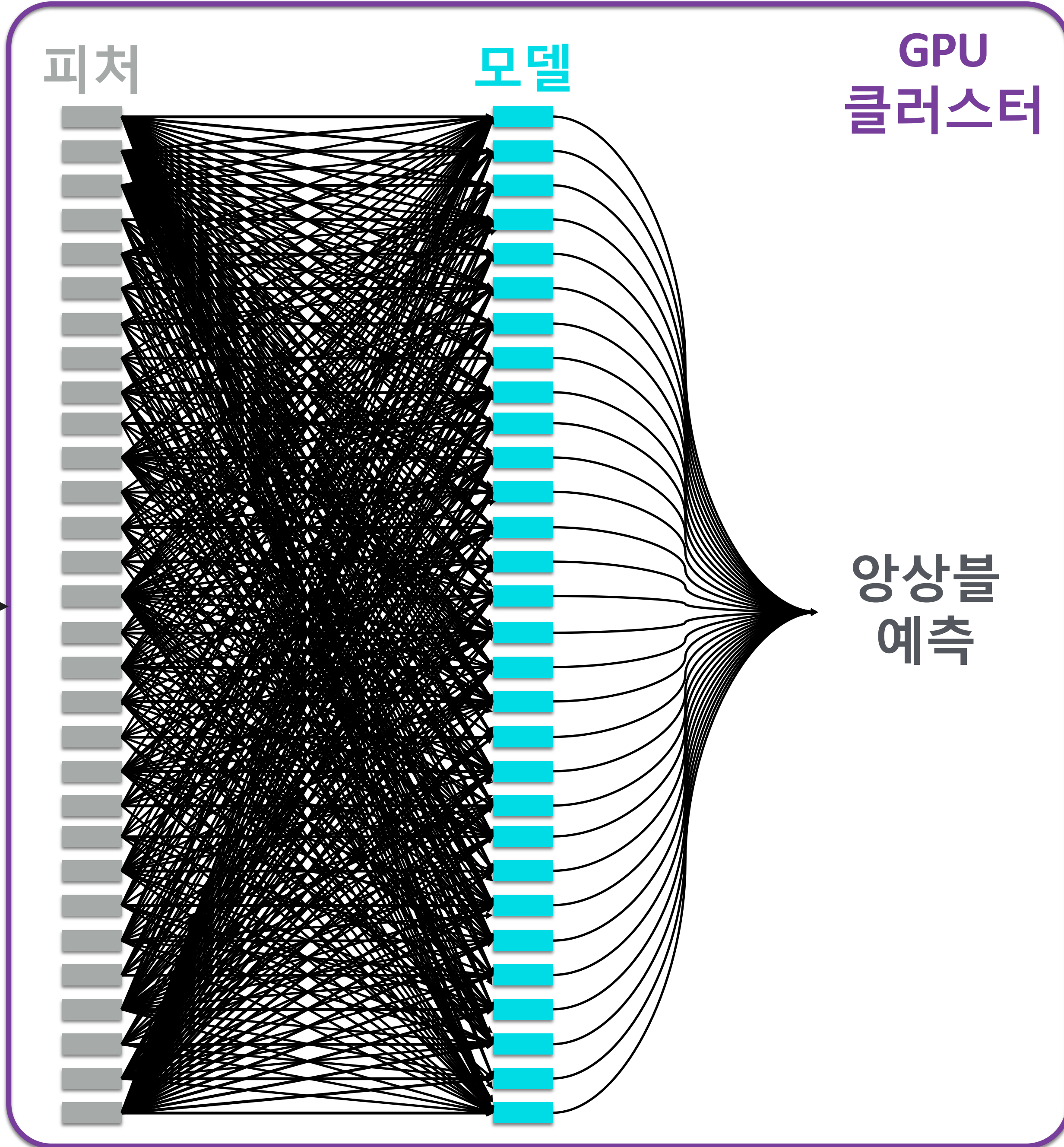
## HyperCLOVA

기존 방식	HyperCLOVA 방식
수동	자동
Supported by human intuition	Supported by gpu cluster & big data
intuitive	Counter intuitive
그럴듯한 피쳐 (ex> 날씨, 광고)	맞춤 피쳐
잘나가는 모델 (Kaggle 1등 모델, public cloud 모델)	맞춤 모델
대략적으로 여러 케이스에 잘 맞춤	내 상황에 잘 맞춤
기성복	맞춤복
소규모 테스트	대규모 테스트

# 1.5 Hyperscale AI



피처제공





## 2. 기존 방식을 사용한 수요예측

# 2.1 네이버 풀필먼트 서비스

news1 | 뉴스1 | 2021.06.22. | 네이버뉴스

## 네이버, AI '클로바 포캐스트'로 물류 혁신..."딥러닝 접목해 기술 고도화"

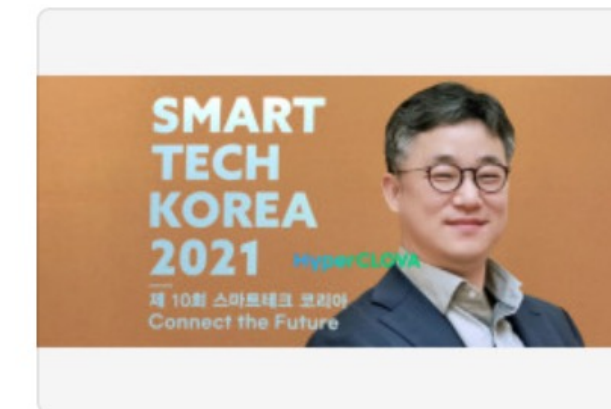
네이버는 클로바 포캐스트를 통해 '오늘 주문-내일 배송' 서비스를 안정적으로 실현하고, 이용자에게 최적의 쇼핑경험을 제공한다는 계획이다. 네이버는 '자사 풀필먼트 서비스를 이용하는 ...



AI타임스 | 2021.06.26.

## 네이버, '클로바 포캐스트'로 물류 혁신한다...마케팅에도 하이퍼클로바 기...

클로바 포캐스트는 지난해 10월부터 네이버 스마트스토어가 대거 입점해 있는 '곤지암 e-풀필먼트 센터'에서 시범 활용 중이다. 물류 AI 사례를 발표 중인 정석근 대표(사진=박성은 기자) ...



TechM | 테크M | 2021.07.20.

## "쿠팡과 다르다" 풀필먼트 플랫폼 띄운 네이버...보안배송에 힘 주는 이유

네이버의 기술력으로 개발한, 물류 수요를 예측하는 인공지능 모델인 '클로바 포캐스트'를 적용, 물류 서비스 효율화를 도모할 예정이다. 또한 올해 중으로 SME와 물류사가 직접 커뮤니케이...



**셀러들의 물품을 물류센터에서 보관부터 포장, 배송까지.  
익일 배송 능력을 강화!**

## 2.2 센터 단위 수요예측이 필요해요!

### 적정 인력 산정

“센터 별 예상 주문 건수에 따라 **적정 패킹 인력**을 미리 산정하고 싶어요.”

### 빅데이터 기반 AI 솔루션

“AI 예측의 도움을 받아 더 **정확하고 자동화된 시스템**을 구축하고 싶어요.”

## 2.3 기존 AI 솔루션들의 성적표

	Model	RMSE	MAPE	Mean NQL[0.1, 0.9]
이동평균	MA	79.28±13.79	0.348±0.015	-
	MLP	<b>72.23±12.60</b>	<b>0.267±0.015</b>	-
	LR	73.20±13.60	0.300±0.010	-
	SARIMA	79.09±17.50	0.318±0.017	0.165±0.033
AWS	DeepAR	76.69±17.86	0.284±0.022	0.132±0.026
	DeepState	79.10±15.83	0.314±0.030	0.152±0.020
Facebook	Prophet	82.21±17.64	0.415±0.018	0.194±0.027

“수요예측 데이터에서는 잘 작동하지 않네?”

## 2.4 무엇이 문제였을까?

판매량  
변동성

부족한  
물류센터  
데이터

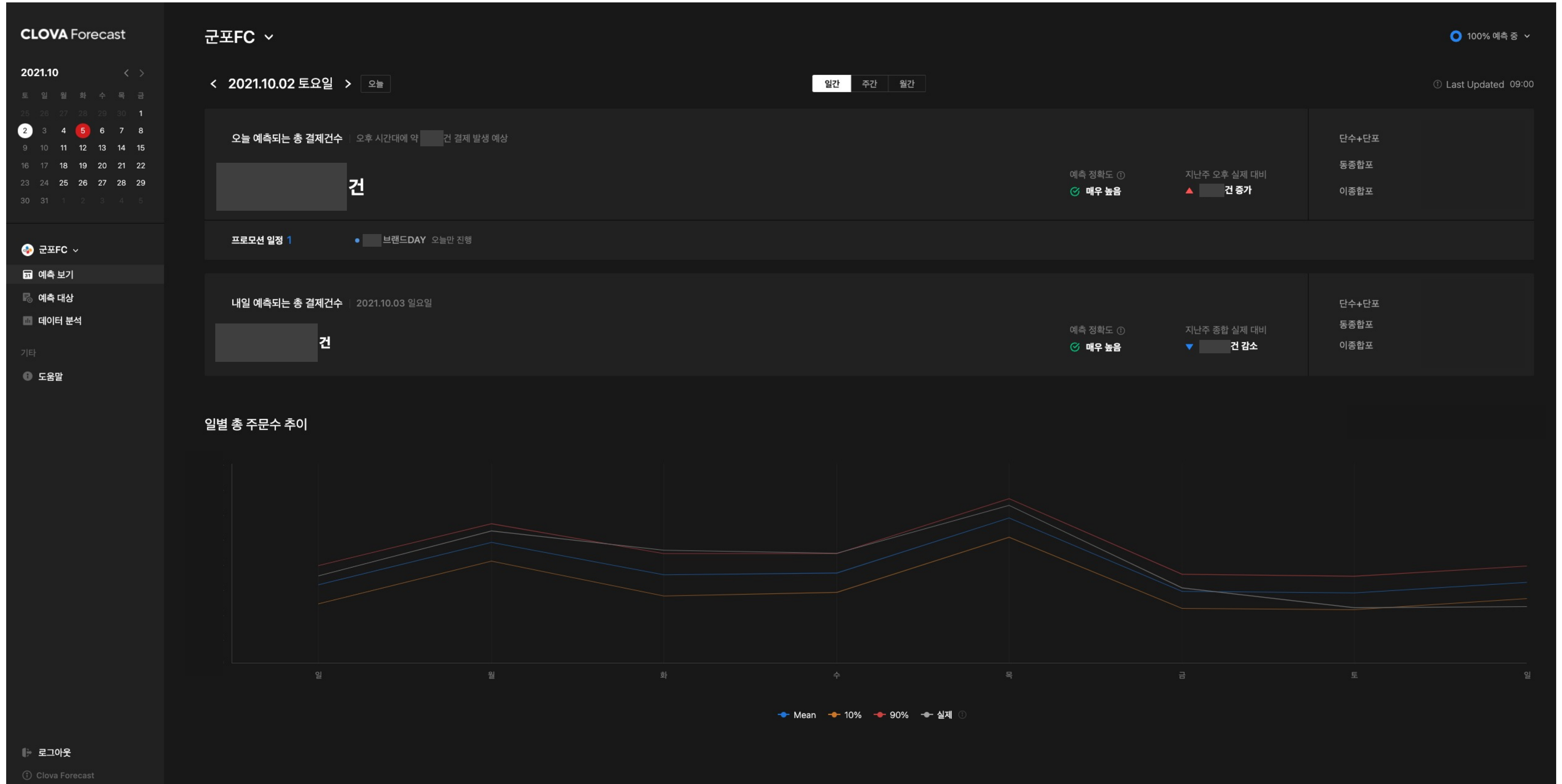
외부 요인  
(프로모션)

1. 꾸준하게 좋은 성적이 나올 수 있는 예측 모델
2. 네이버 데이터를 잘 활용할 수 있는 예측 모델

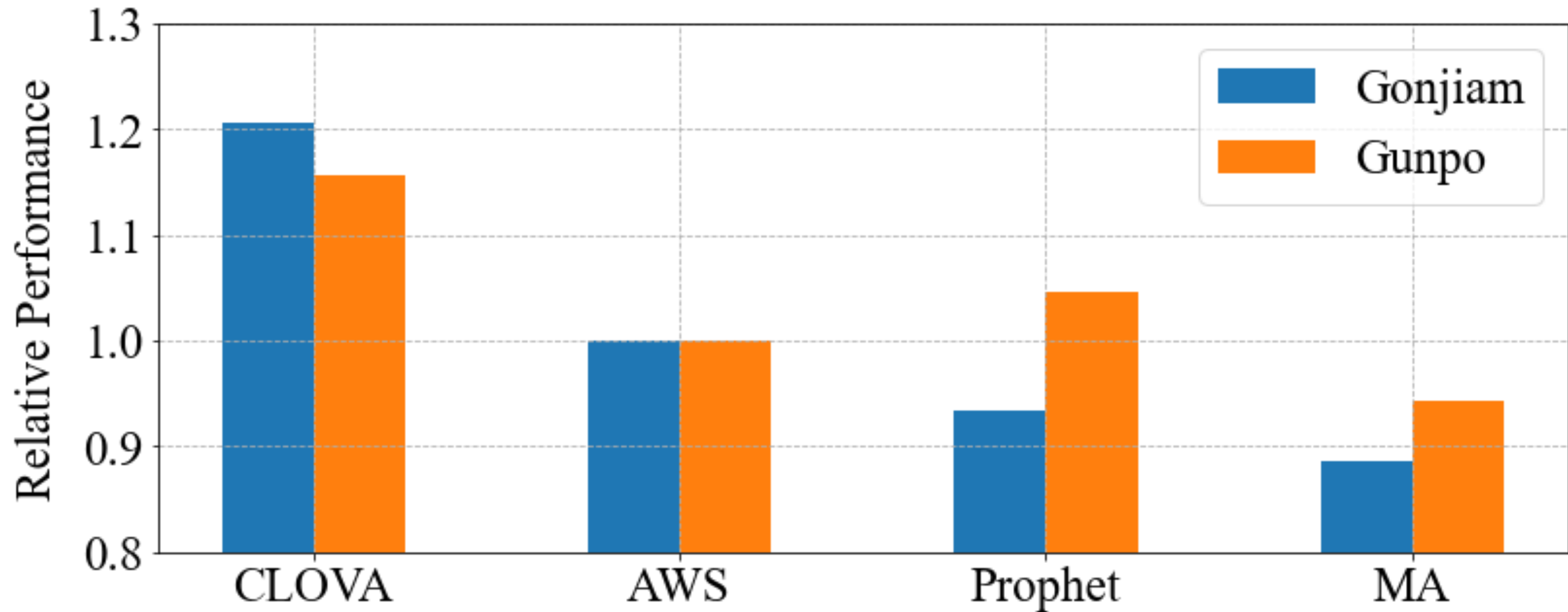
## 2.5 클로바 포캐스트 (Workflow)



# 2.6 클로바 포캐스트 (Dashboard)



## 2.7 정확도는 높은가?

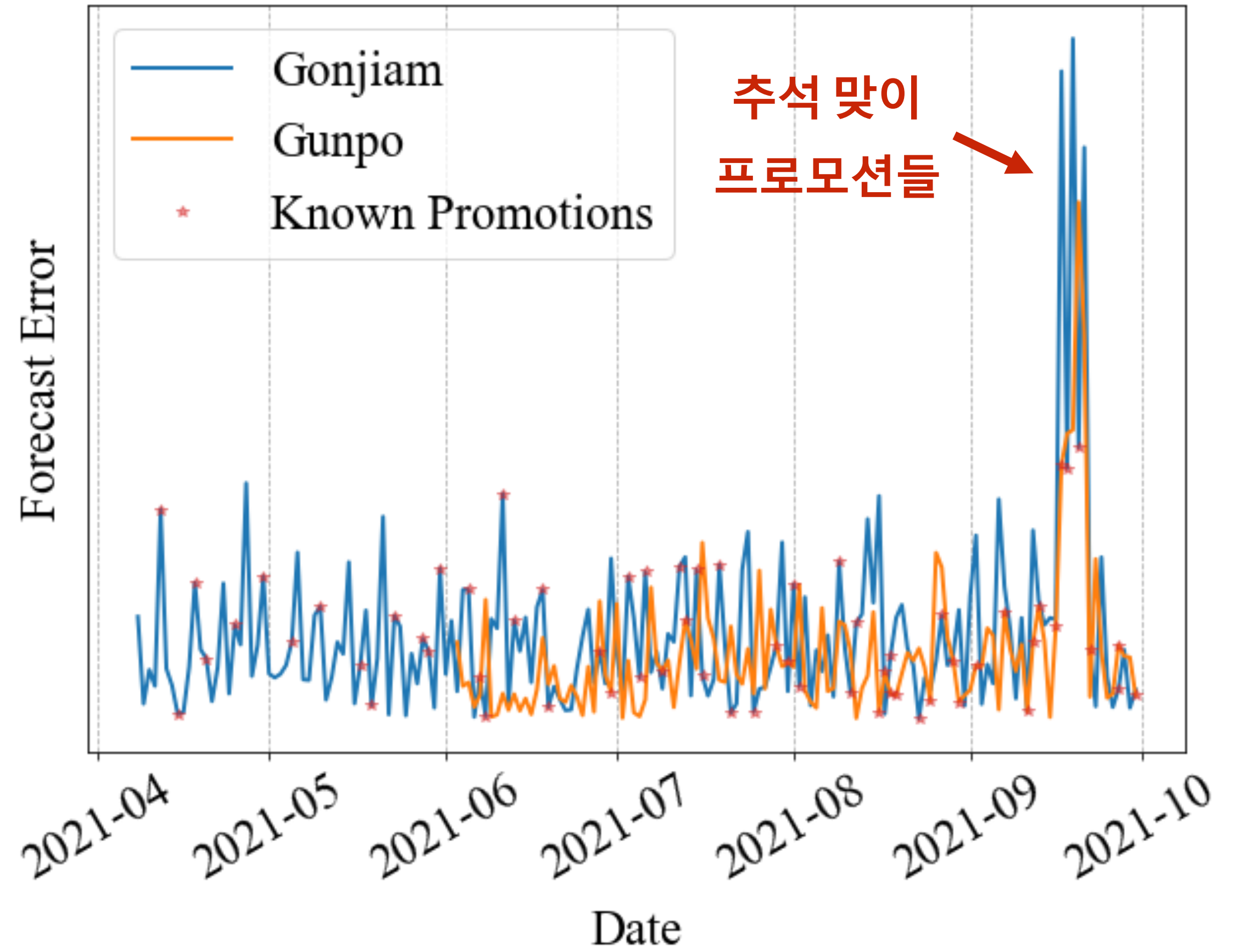
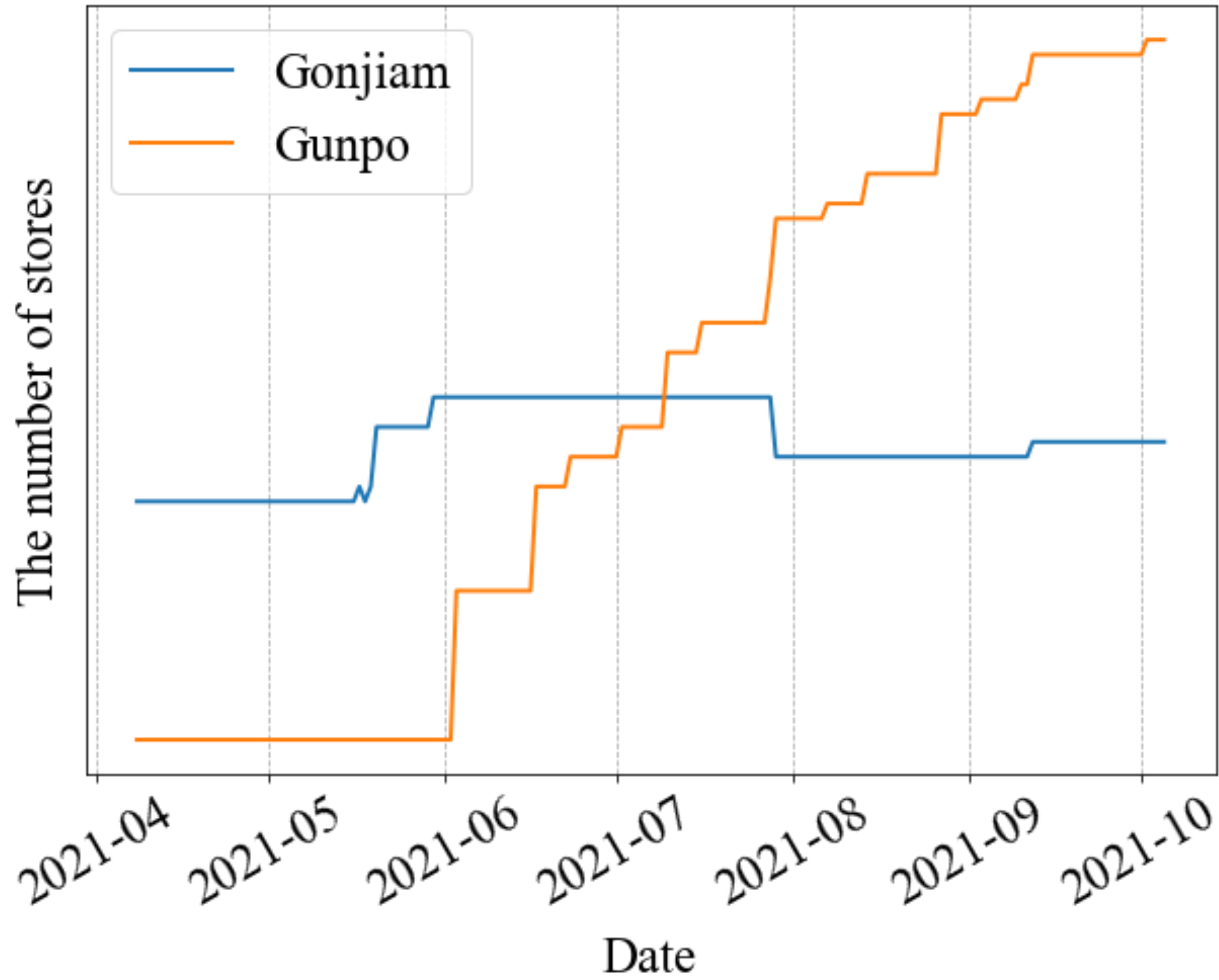


AWS model is implemented by GluonTS.

MA = Moving Average.

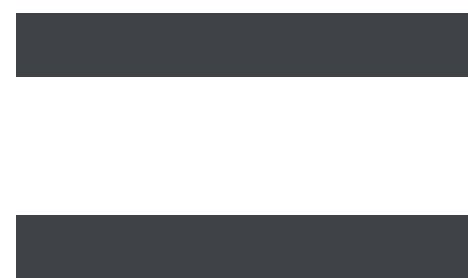


# 2.8 일정한 성능이 나오는가?



# 2.9 SKU 단위 수요예측도 필요해요!

- SKU (Stock Keeping Unit): 상품 재고 관리 단위



LG생... ₩825,809  찜한스토어  소식받는다

온더바디 퍼퓸 러브, 퍼퓸 바디워시

[구매하기](#) [장바구니 담기](#) [선물하기](#)

네이버 쇼핑 상품

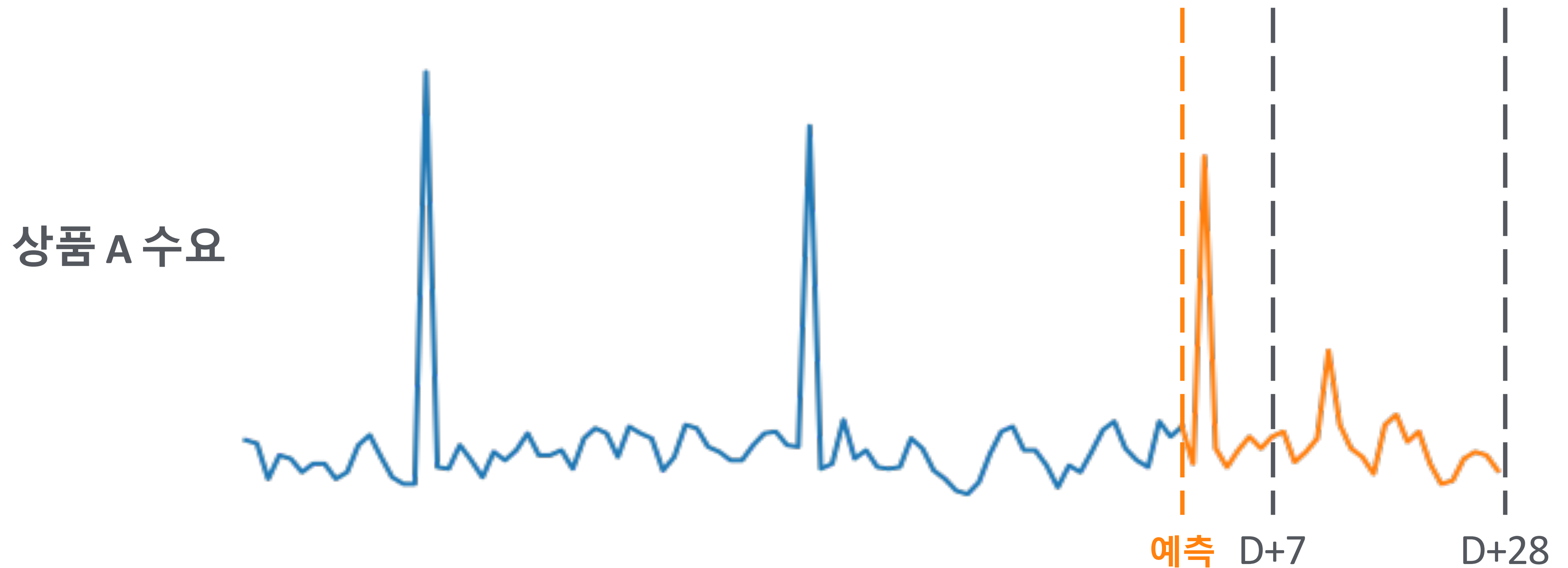
SKU 1

(8801051130690)

SKU 2

(8801051130683)

## 2.9 SKU 단위 수요예측도 필요해요!



“효율적인 상품 재고 관리를 통해 악성 재고를 줄이고, 품절에 대비하고 싶어요.”

## 2.10 왜 예측하기 어려운가?

Sparse한  
시계열  
데이터

SKU별  
각기 다른  
시계열 특성

수 많은  
(대부분 쓸모 없는)  
피처

## 2.11 다양한 모델 x 다양한 학습 방법

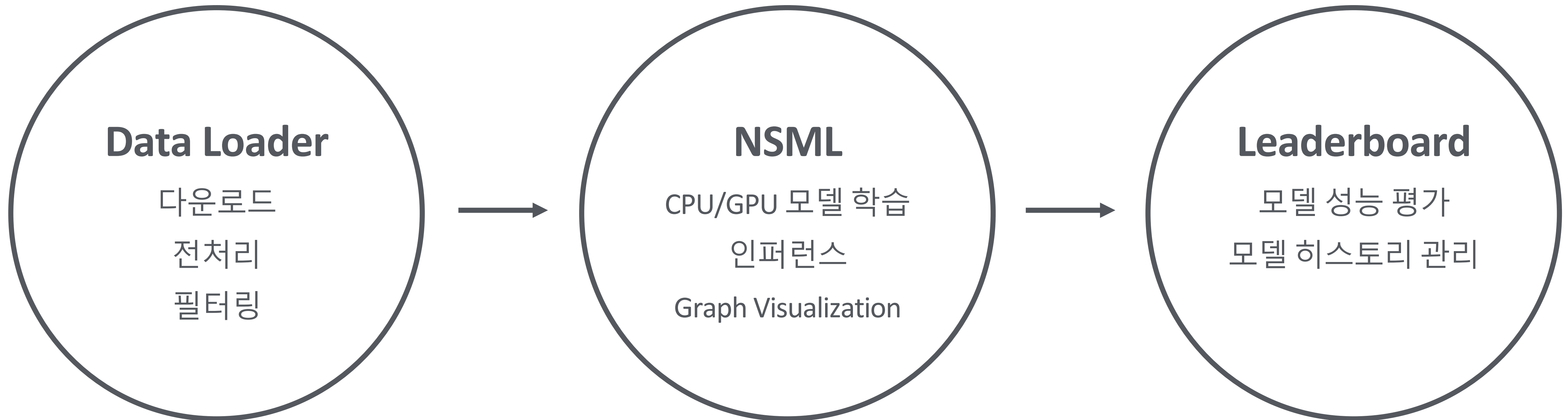
### 모델 (Model Architectures)

- Classical: Moving Average, ARIMA,
- Amazon Forecast: Prophet, CNN-QR, DeepAR, DeepState
- CLOVA Forecast: Linear, MLP, LSTM, Transformer, Multivariate Forecasting, ...

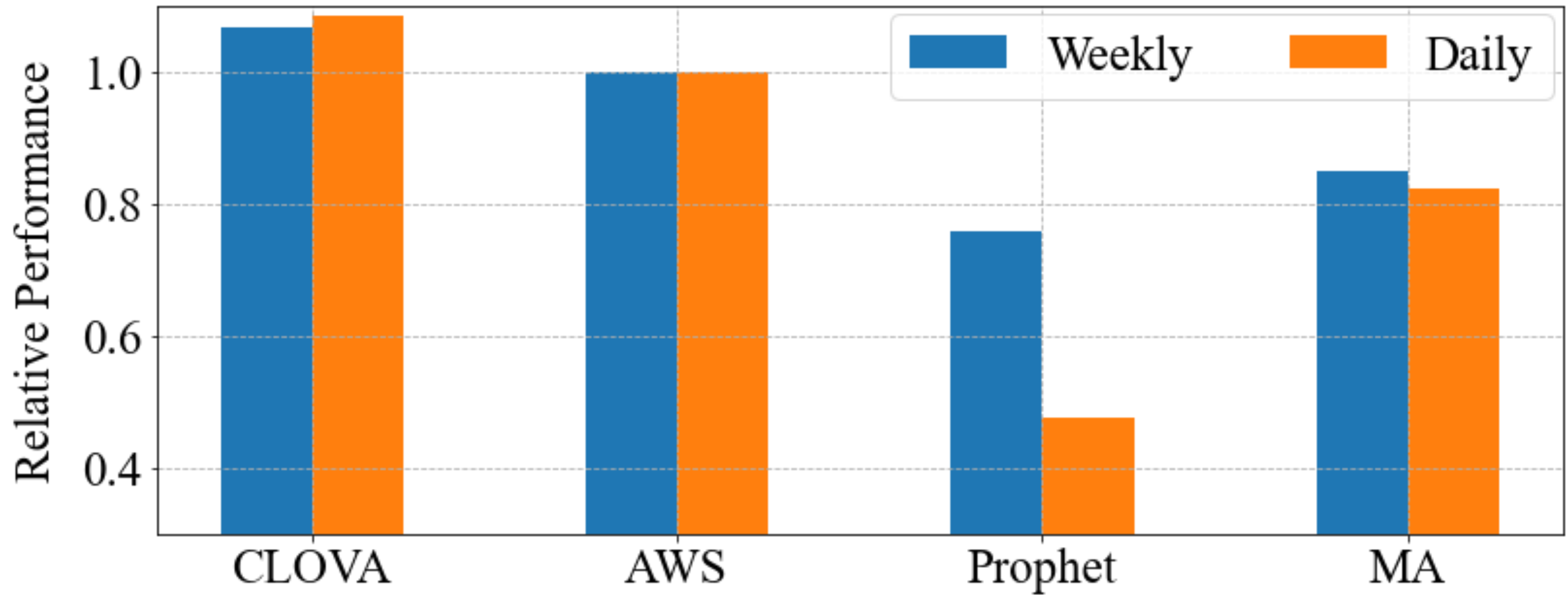
### 학습 방법 (Training Strategies)

- Output Distributions & Loss functions
- Normalization & Augmentation
- Sampling Strategy
- Regularization

## 2.12 모델 테스트 환경



# 2.13 예측 정확도



AWS model is implemented by GluonTS.

MA = Moving Average.

## 2.14 클로바가 얻은 교훈

Q. 더 잘하기 위해서는?

- 다양한 피처 (네이버 데이터) 테스트 & 활용 (Ch.3)
- 각각의 모델의 장점을 살릴 수 있는 방법 (Ch.4)

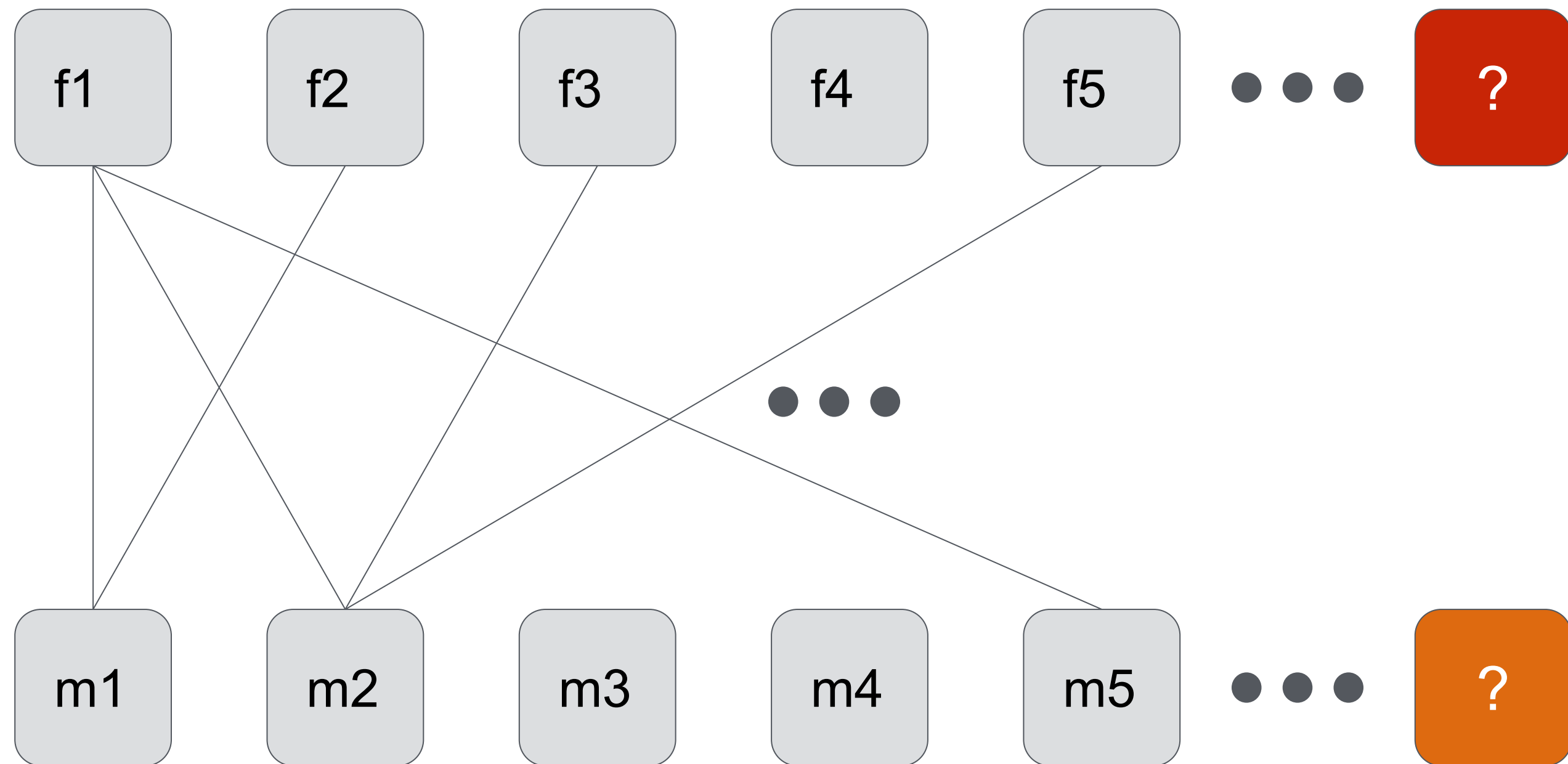
→ **HyperCLOVA**



# 3. HyperCLOVA 수요예측

# 3. HyperCLOVA 수요예측: 문제

Feature Pool



Model Pool

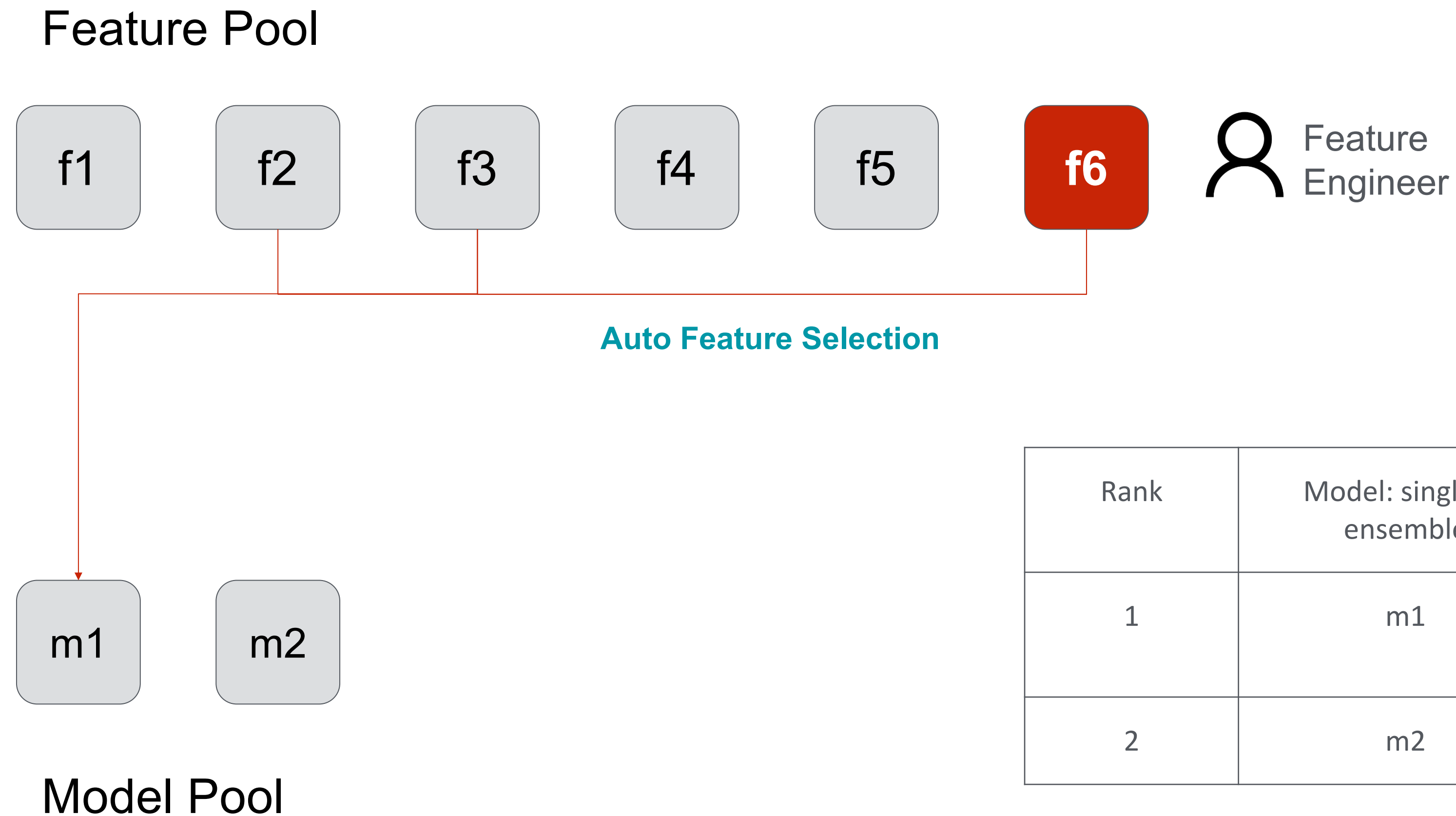
Feature 조합 경우의 수:  $\sum_i^n i C_n$

- FC(5): 26
- FC(10): 1,013
- FC(20): 1,048,555

Hyperparameter 조합수:  $10^3$

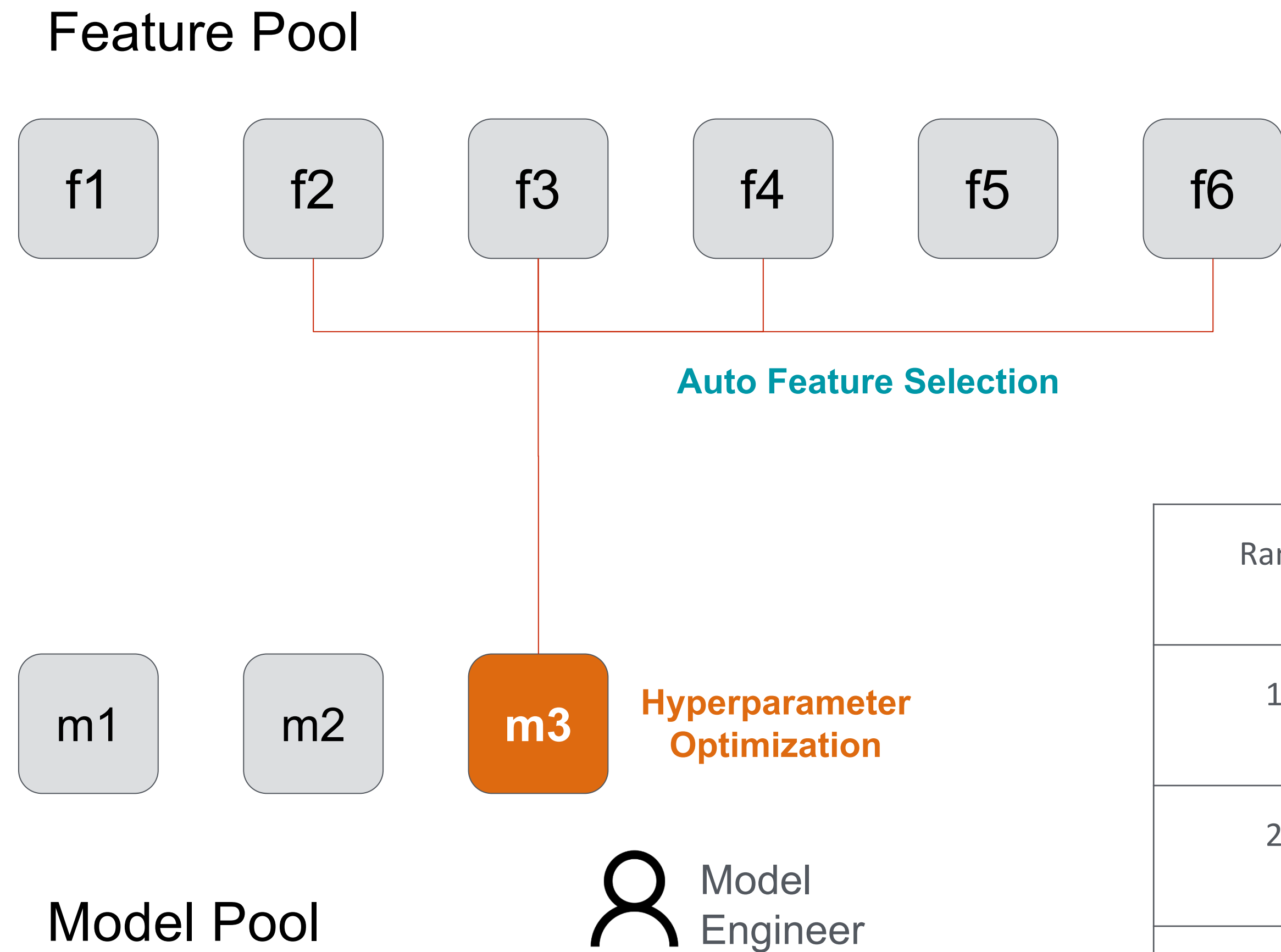
- Parameter: 3, 10개 Grid 서치

# 3. HyperCLOVA 수요예측: 파이프라인



Rank	Model: single or ensemble	Feature Combination	Score	Contributor
1	m1	f2,f3, <b>f6</b>	0.23	Feature engineer
2	m2	f1,f2,f5	0.3	

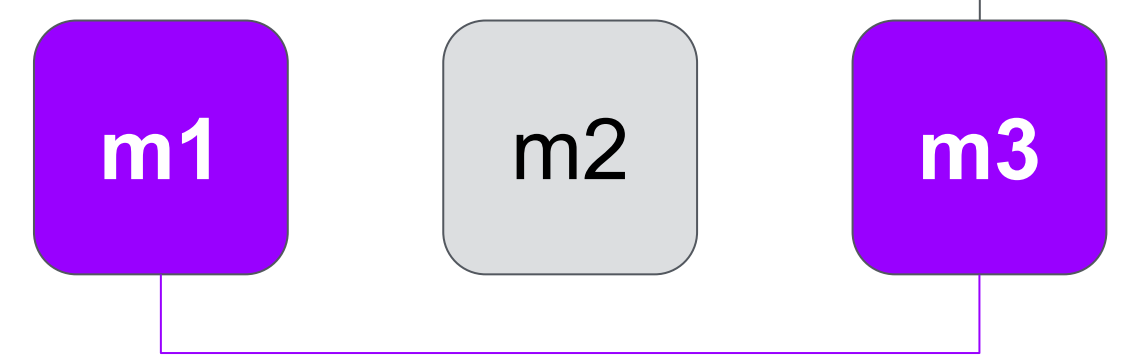
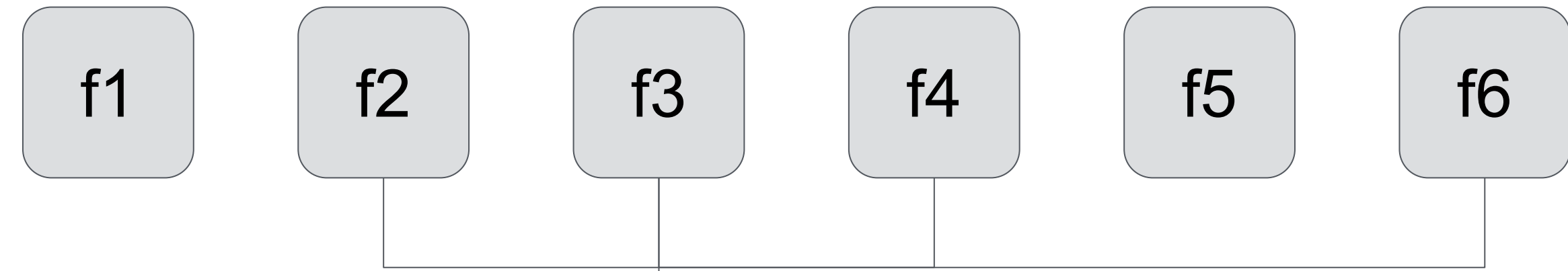
# 3. HyperCLOVA 수요예측: 파이프라인



Rank	Model: single or ensemble	Feature Combination	Score	Contributor
1	m1	f2,f3,f6	0.23	Feature engineer
2	<b>m3</b>	f2,f3,f4,f6	0.25	Model engineer
3	m2	f1,f2,f5	0.3	

# 3. HyperCLOVA 수요예측: 파이프라인

Feature Pool

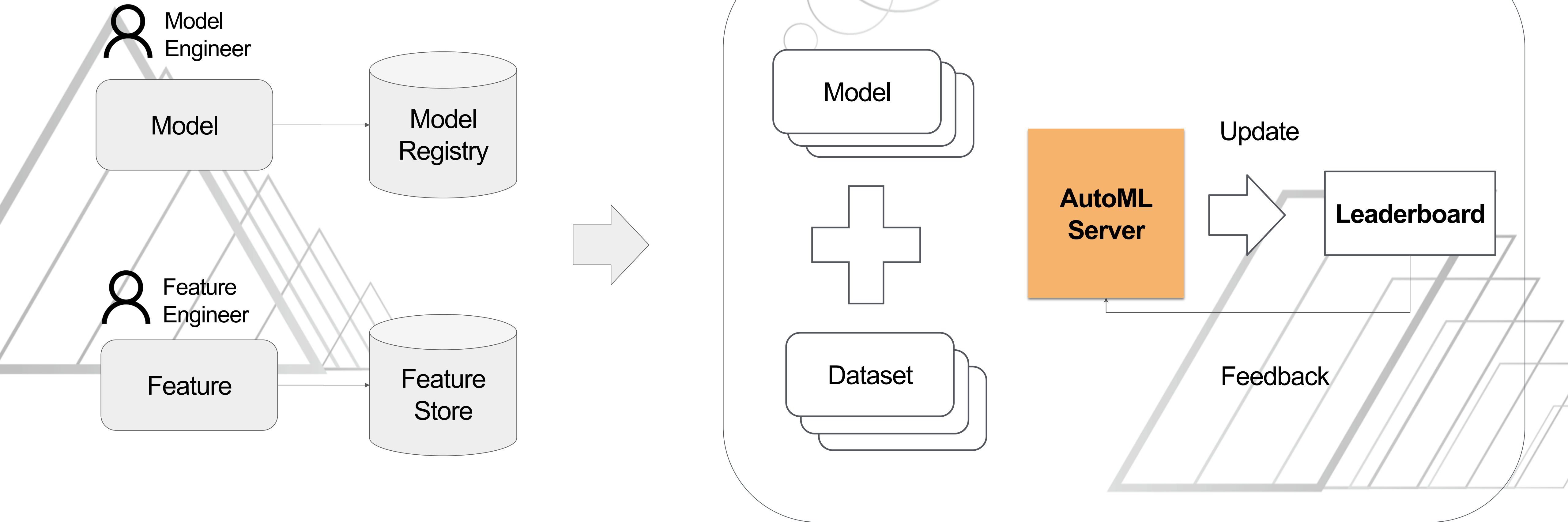


Ensemble Methods

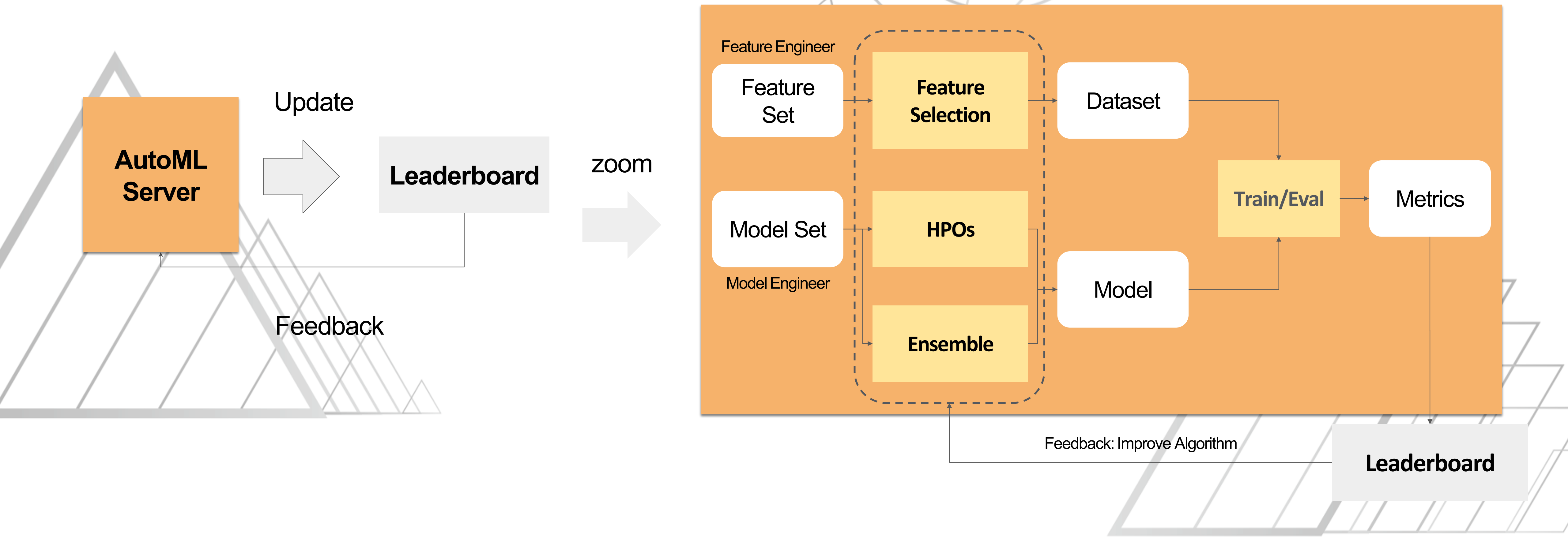
Model Pool

Rank	Model: single or ensemble	Feature Combination	Score	Contributor
1	en(m1, <b>m3</b> )	f2,f3,f4, <b>f6</b>	0.2	Model engineer Feature engineer
2	m1	f2,f3, <b>f6</b>	0.23	Feature engineer
3	<b>m3</b>	f2,f3,f4, <b>f6</b>	0.25	Model engineer
4	m2	f1,f2,f5	0.3	

### 3. HyperCLOVA 수요예측: 파이프라인



# 3. HyperCLOVA 수요예측: 파이프라인



# 4. HyperCLOVA 활용 결과 (SKU단위 수요예측)



# 4.1 “All models are wrong.” - George E.P. Box

SKU 수요예측 문제에서 모델 채택에 관한 문제:

- 내일/다음주/다음달 예측을 위해 어떤 모델의 결과를 내보내야 할까?
- 모델 리더보드는 모델 들의 과거 특정 기간 (test period) “평균” 퍼포먼스만을 알려준다.
- 과거의 최적 평균 퍼포먼스 =? 미래의 최적 퍼포먼스
- 오버피팅은?
- SOTA 모델 이외의 다른 모델들은 쓸모가 없는가?

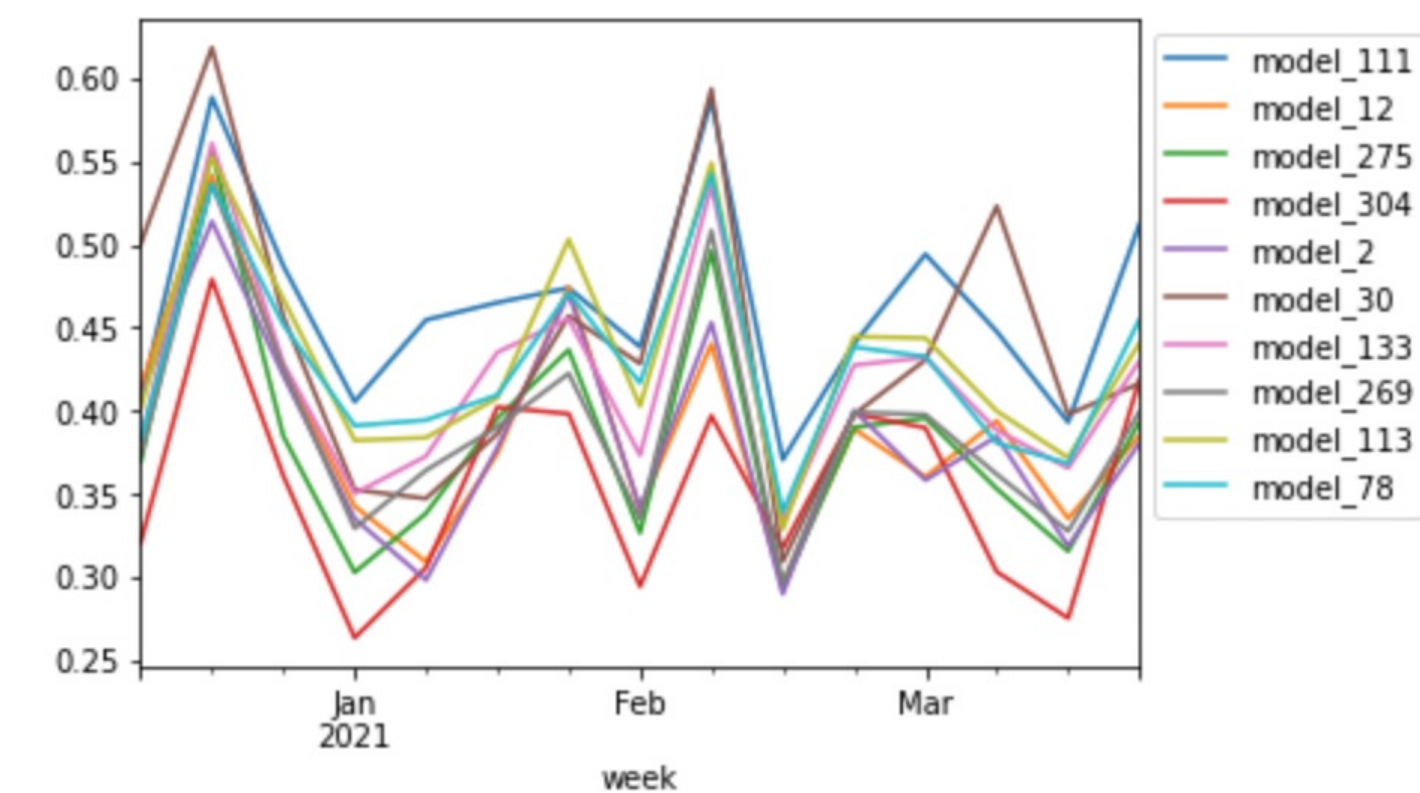
## 2.11 다양한 모델 x 다양한 학습 방법

모델 (Model Architectures)

Classical: Moving Average, ARIMA,  
Amazon Forecast: Prophet, CNN-QR, DeepAR, DeepState  
CLOVA Forecast: Linear, MLP, LSTM, Transformer, Multivariate Forecasting, ...

학습 방법 (Training Strategies)

Output Distributions & Loss functions  
Normalization & Augmentation  
Sampling Strategy  
Regularization

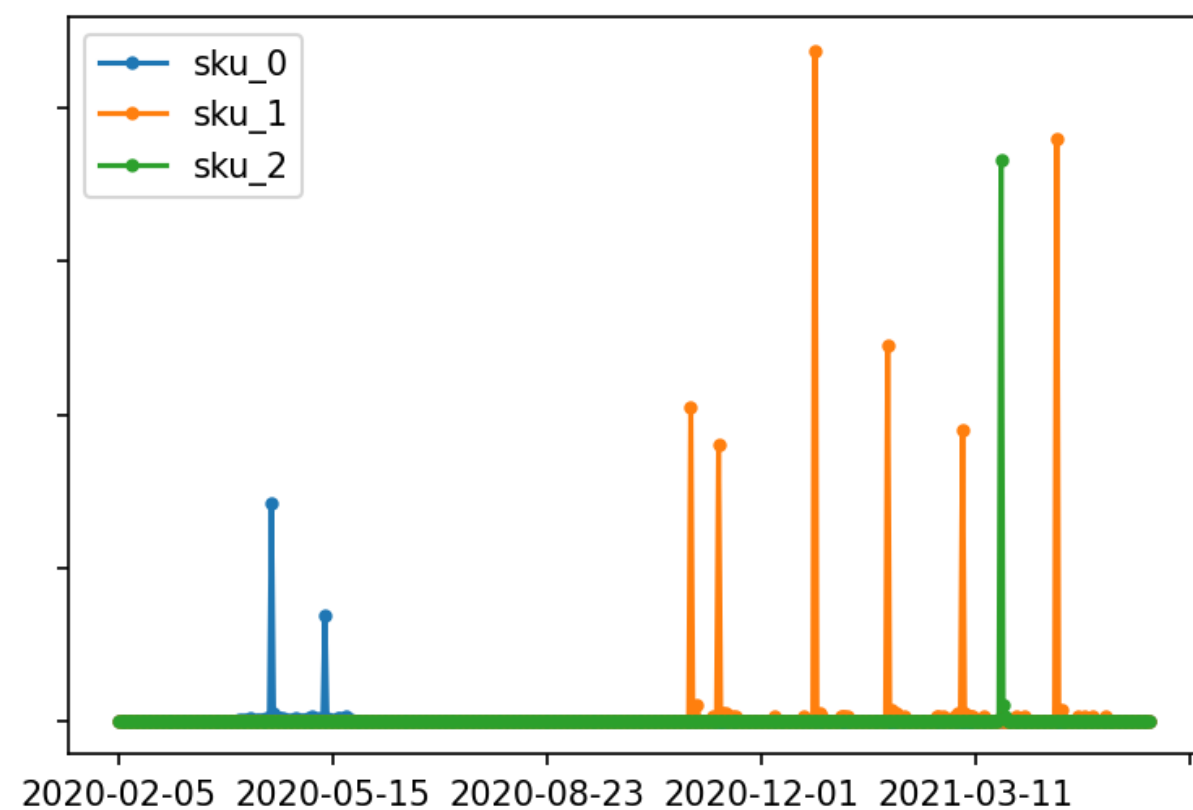


# 4.2 The dimensionality of the problem is huge

SKU space is of million scale

	sku_1	sku_3	sku_4	sku_5	sku_6	sku_7	sku_8	sku_9	sku_10	sku_11	...	s
2021-02-21	0.0	40.0	132.0	0.0	0.0	0.0	106.0	0.0	0.0	0.0	...	
2021-02-22	0.0	60.0	140.0	0.0	0.0	0.0	112.0	0.0	0.0	0.0	...	
2021-02-23	1.0	32.0	232.0	0.0	0.0	0.0	82.0	0.0	0.0	0.0	...	
2021-02-24	0.0	58.0	154.0	0.0	0.0	0.0	96.0	0.0	0.0	0.0	...	
2021-02-25	0.0	16.0	154.0	20.0	0.0	0.0	68.0	14.0	0.0	0.0	...	
...	...	...	...	...	...	...	...	...	...	...	...	...

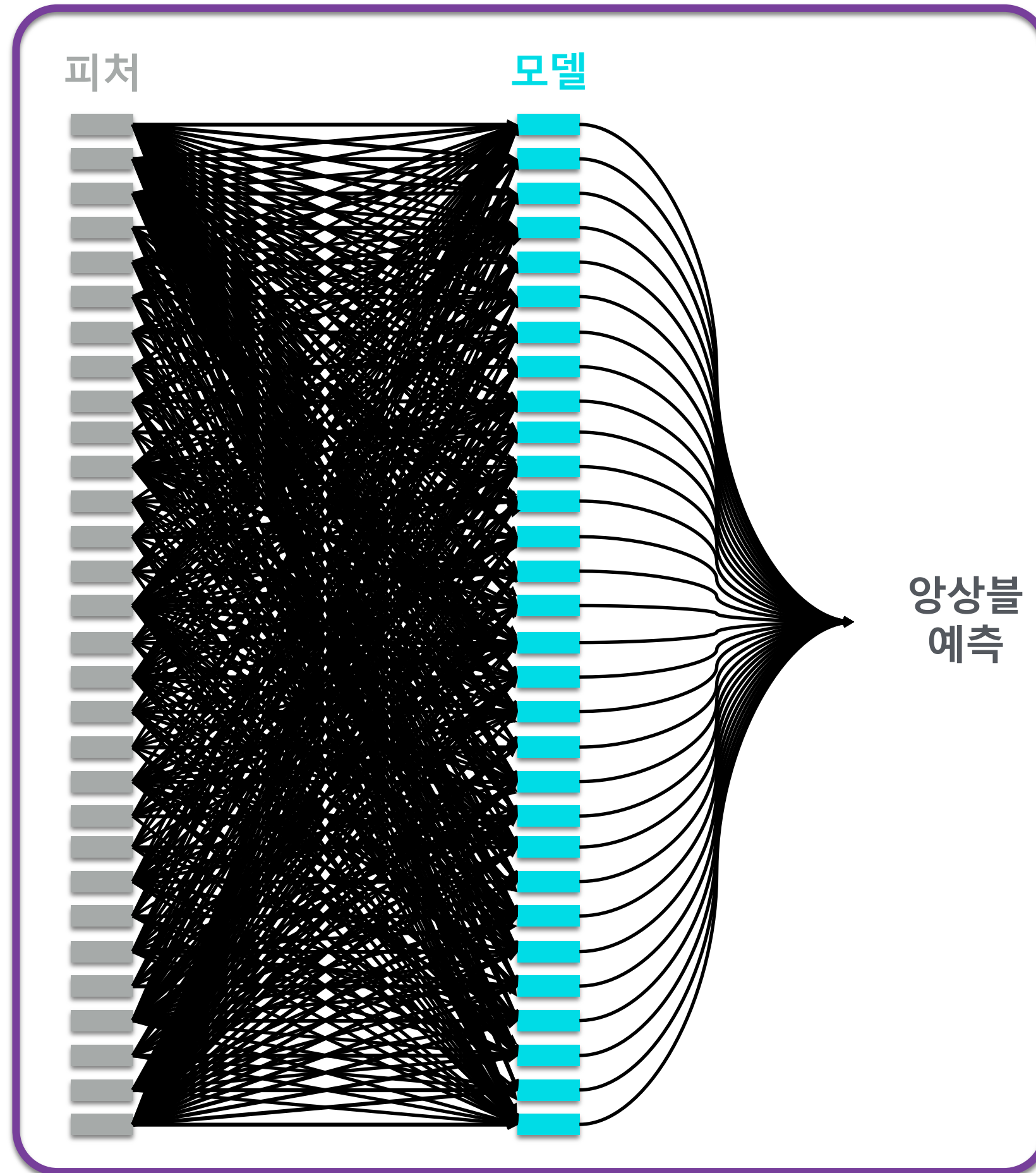
and sparse...



## Large dimension problem

- SKU time-series of million scale
- Different stores, categories, ...
- Seasonality
- event driven anomalies
- ...
- 데이터는 부족 (sparse, not so long historical period)
- 안정적이고 견고한 개별 모델을 만드는게 쉽지 않음
- 앙상블 (Ensemble) 접근이 합리적

## 4.3 Ensemble approach



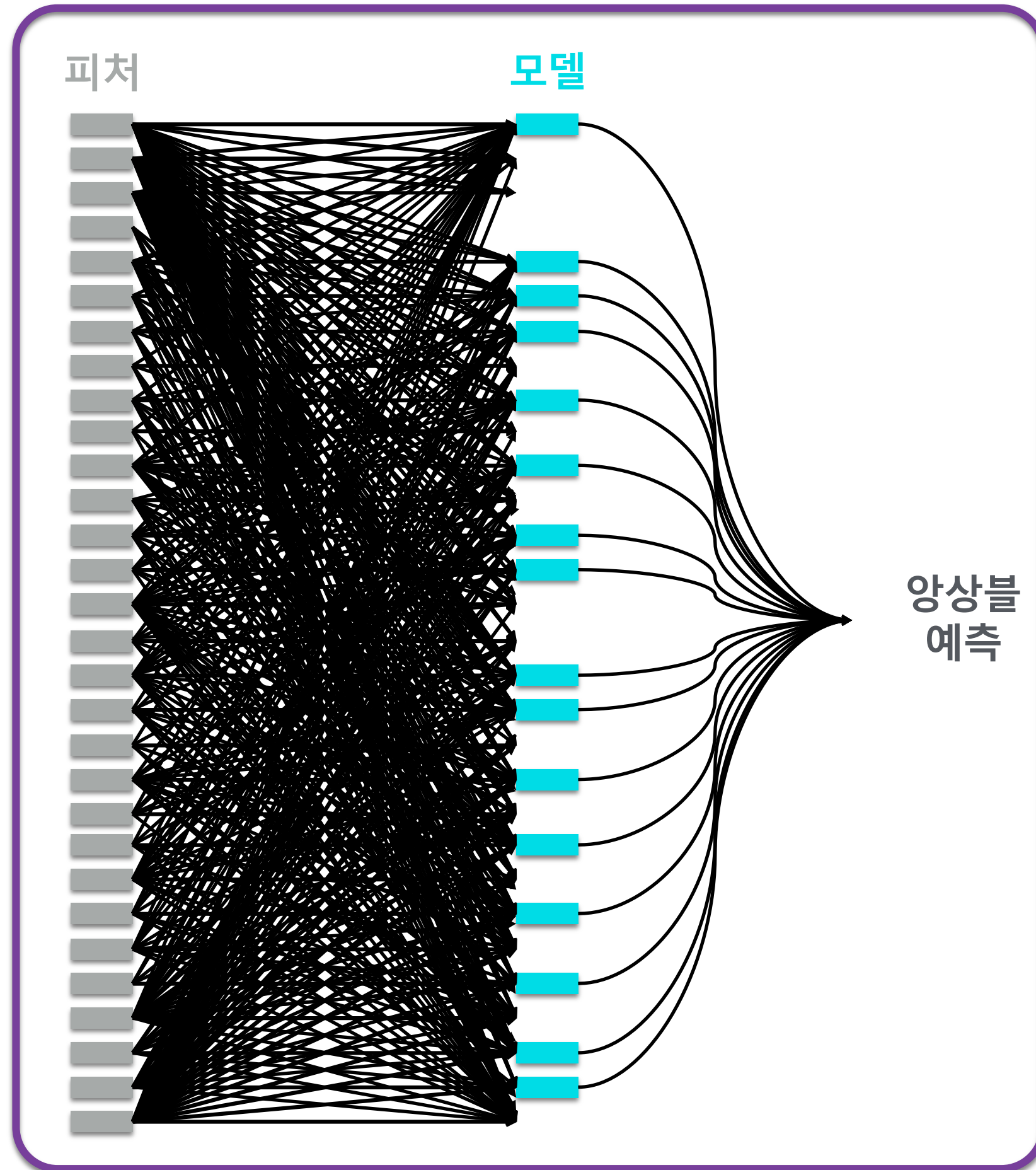
개별 모델의 예측치: 새로운 데이터셋

$$\hat{y} = \sum_k w_k M_k(X)$$

최적화 문제 (가중치 결정):

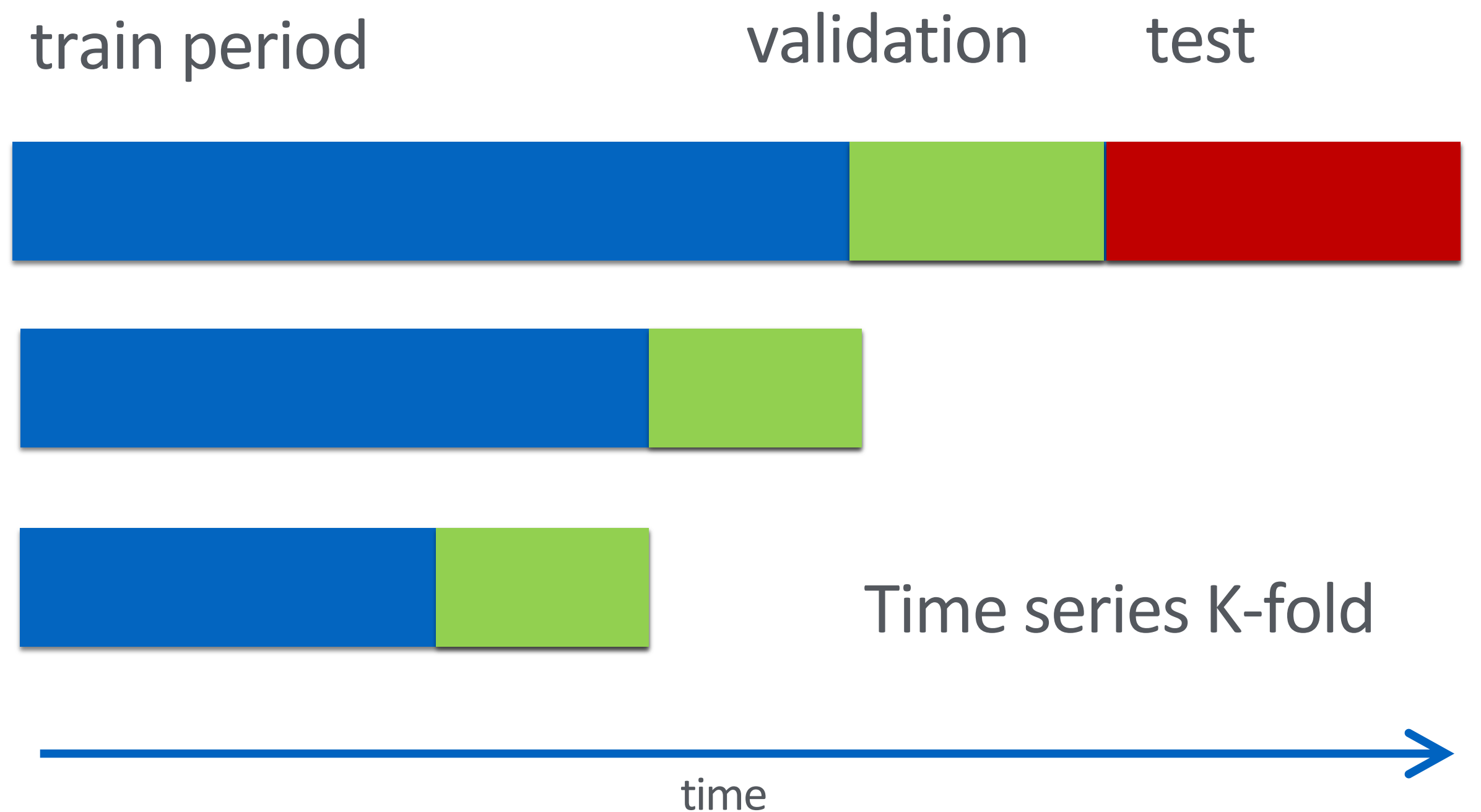
$$\operatorname{argmin}_{\{w\}} (\operatorname{Metric}(y, \hat{y}\{w\}))$$

# 4.4 Hyperparameter space is huge

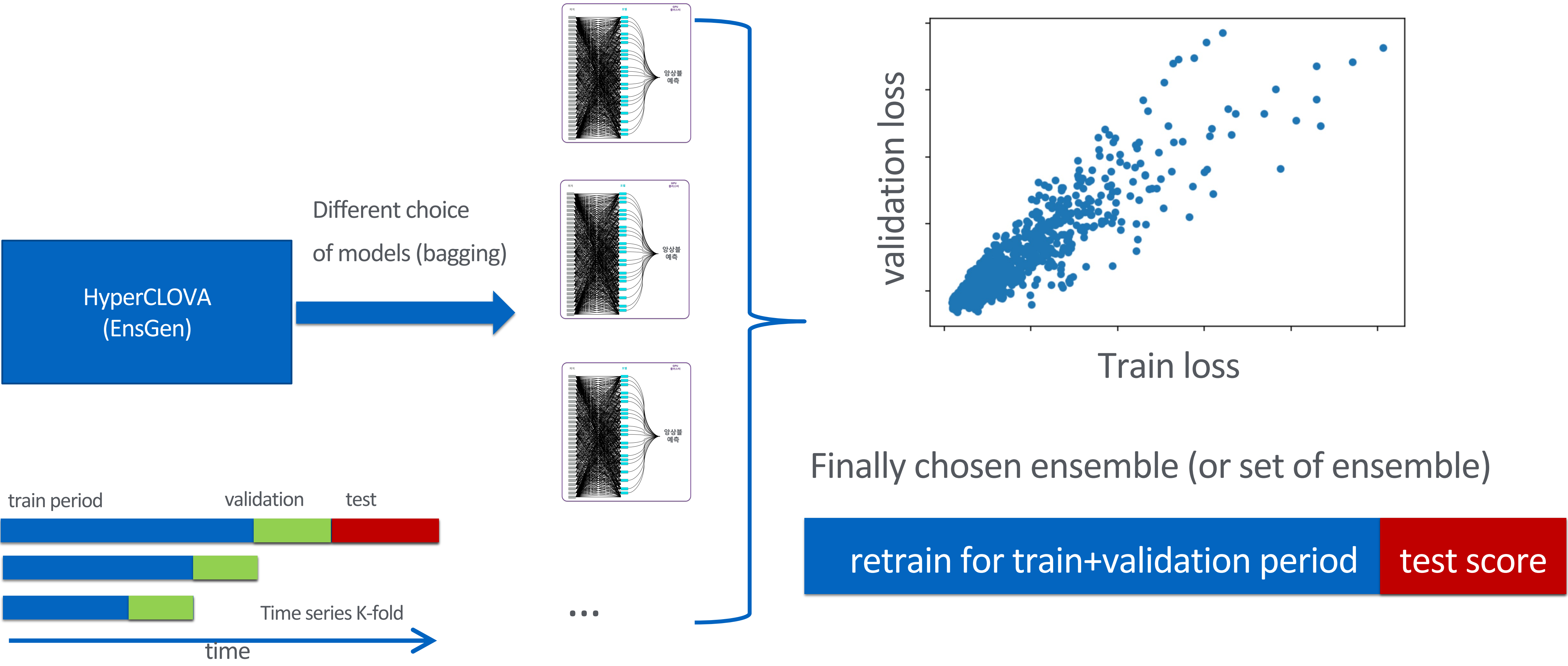


Hyper parameter: choice of models

- N: total model pool
- p: number of chosen model set
- >  $k * {}_N C_p$  number to be tested..

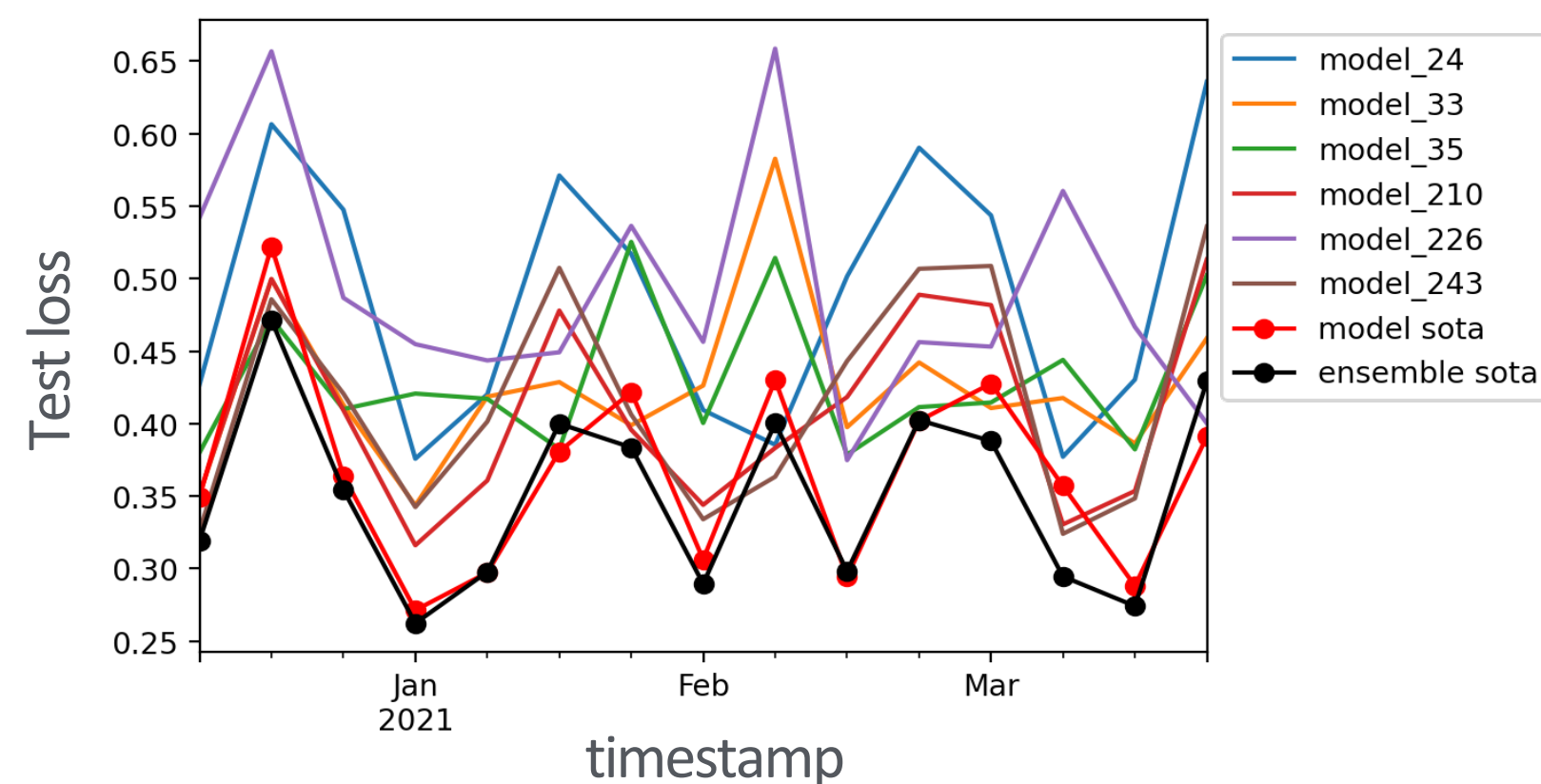


# 4.5 HyperCLOVA: ensgen (Ensemble Generator package)



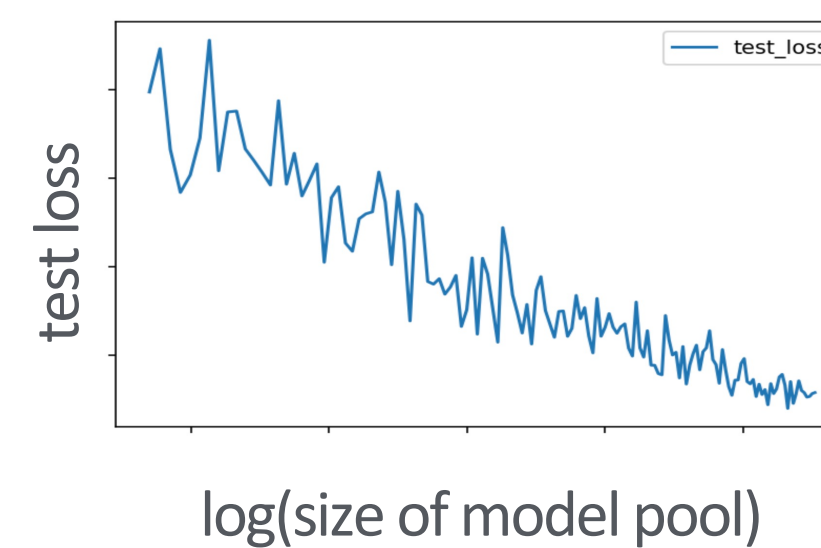
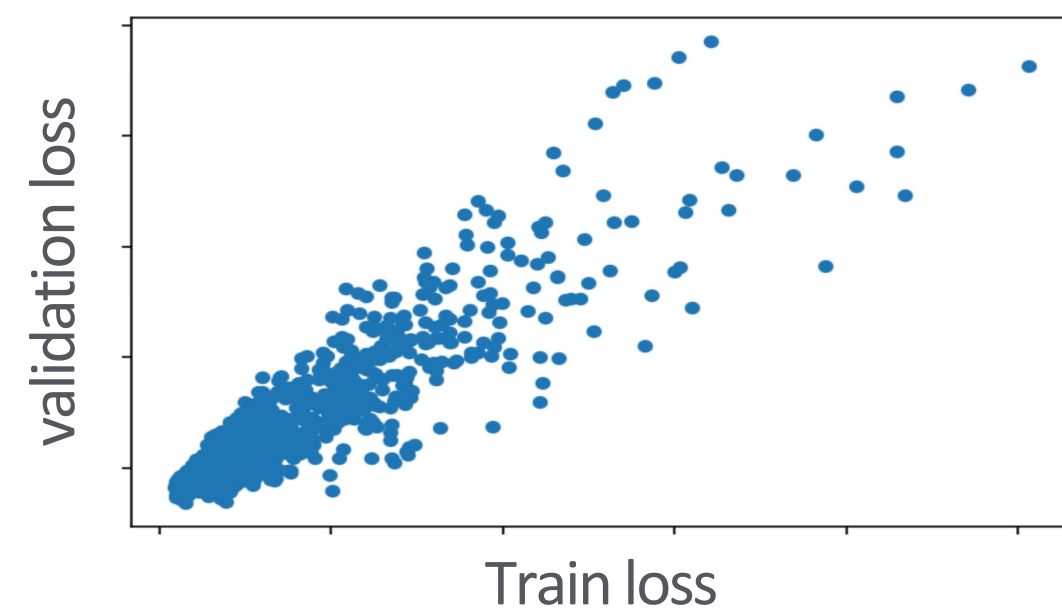
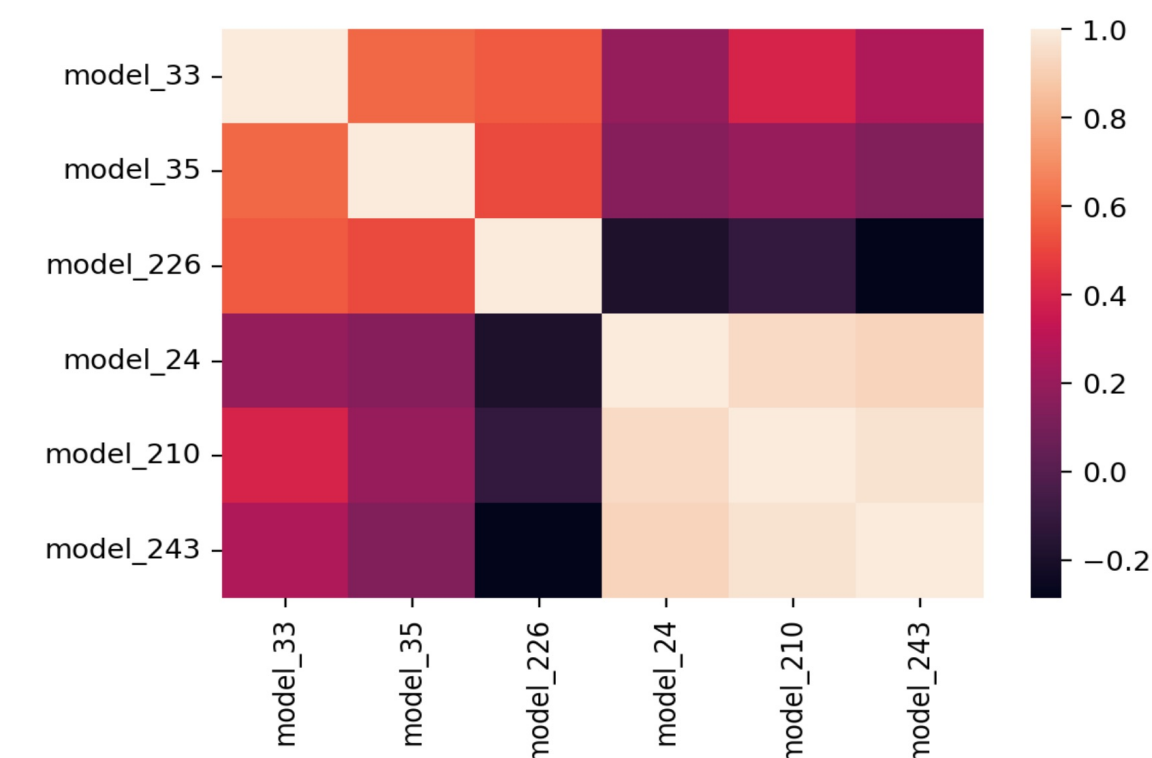
# 4.6 Ensemble result

Top weighted models in the ensemble



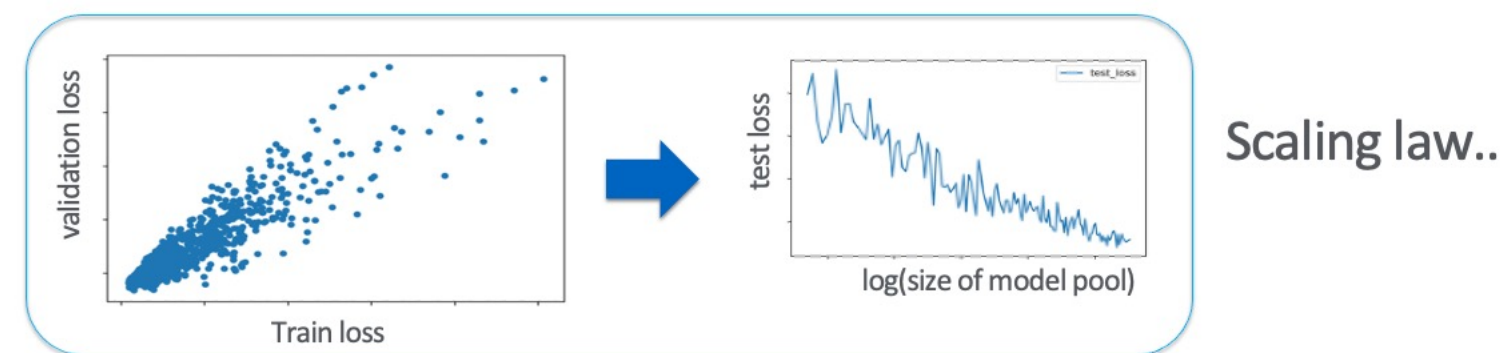
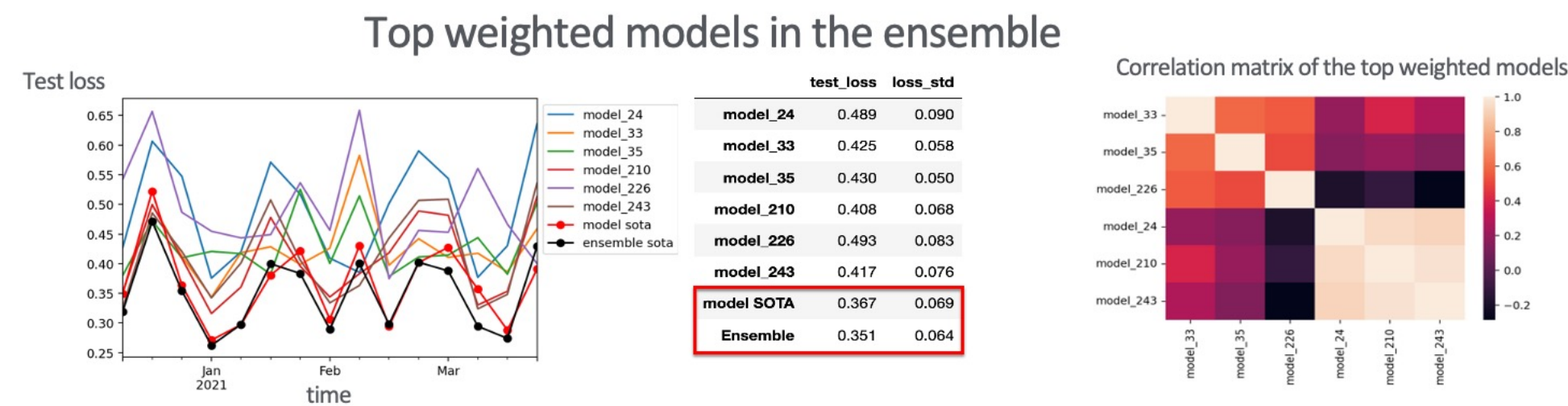
	test_loss	loss_std
<b>model_24</b>	0.489	0.090
<b>model_33</b>	0.425	0.058
<b>model_35</b>	0.430	0.050
<b>model_210</b>	0.408	0.068
<b>model_226</b>	0.493	0.083
<b>model_243</b>	0.417	0.076
<b>model SOTA</b>	0.367	0.069
<b>Ensemble</b>	0.351	0.064

Correlation matrix of the top weighted models



Scaling law..

# 4.7 Impact of HyperCLOVA



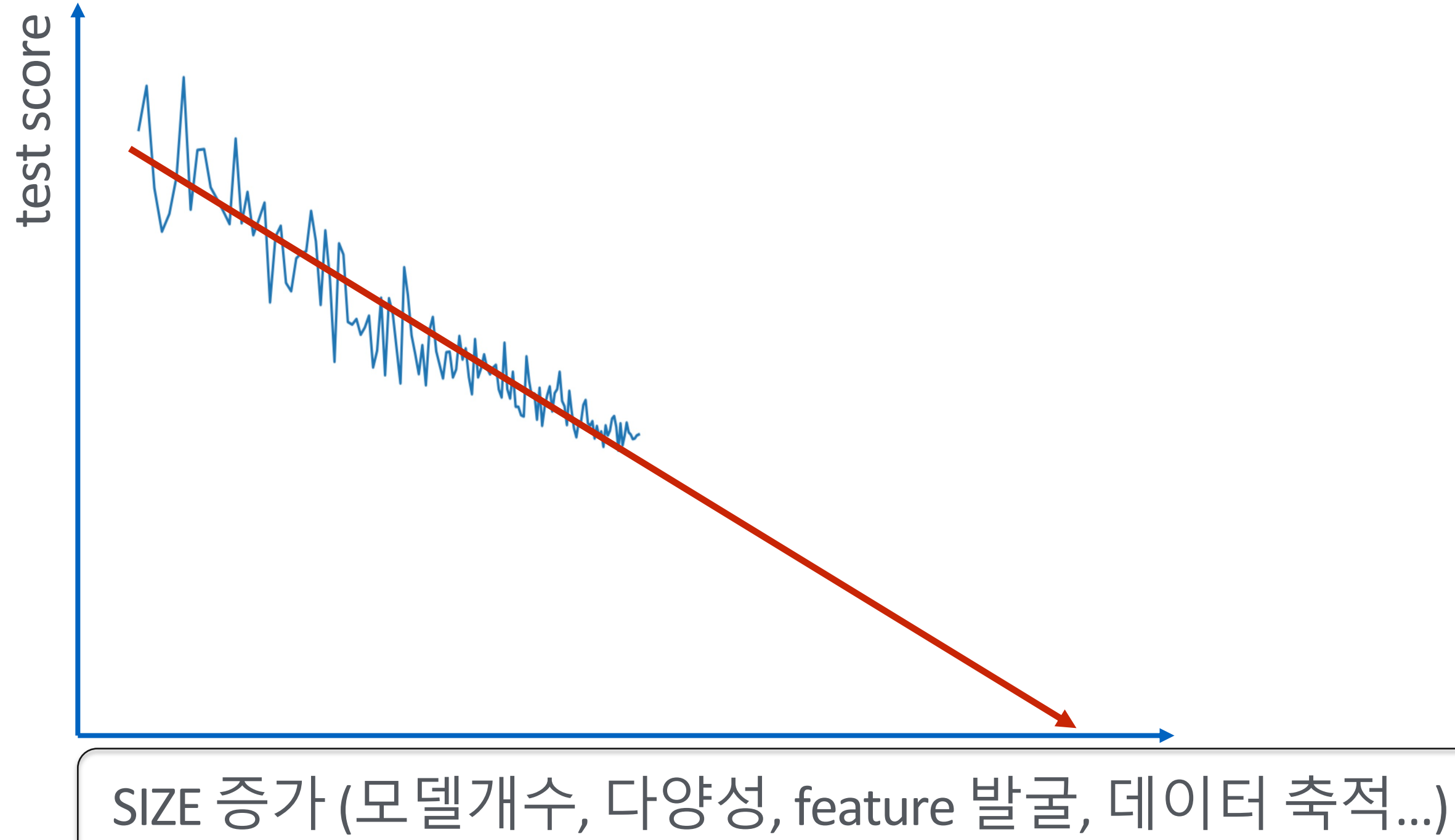
- HyperCLOVA framework: **Human intuition과 반대의 결과를 얻음** -> 성능이 좋은 모델을 모아야 더 좋은 결과가 달성되는 것이 아님.

- HyperCLOVA scheme에서는 SOTA 모델을 달성하기 위한 노력이 필요 없음. 개별 스코어와 관계없이 서로 다른 dimension을 타겟하는 모델들이 필요.

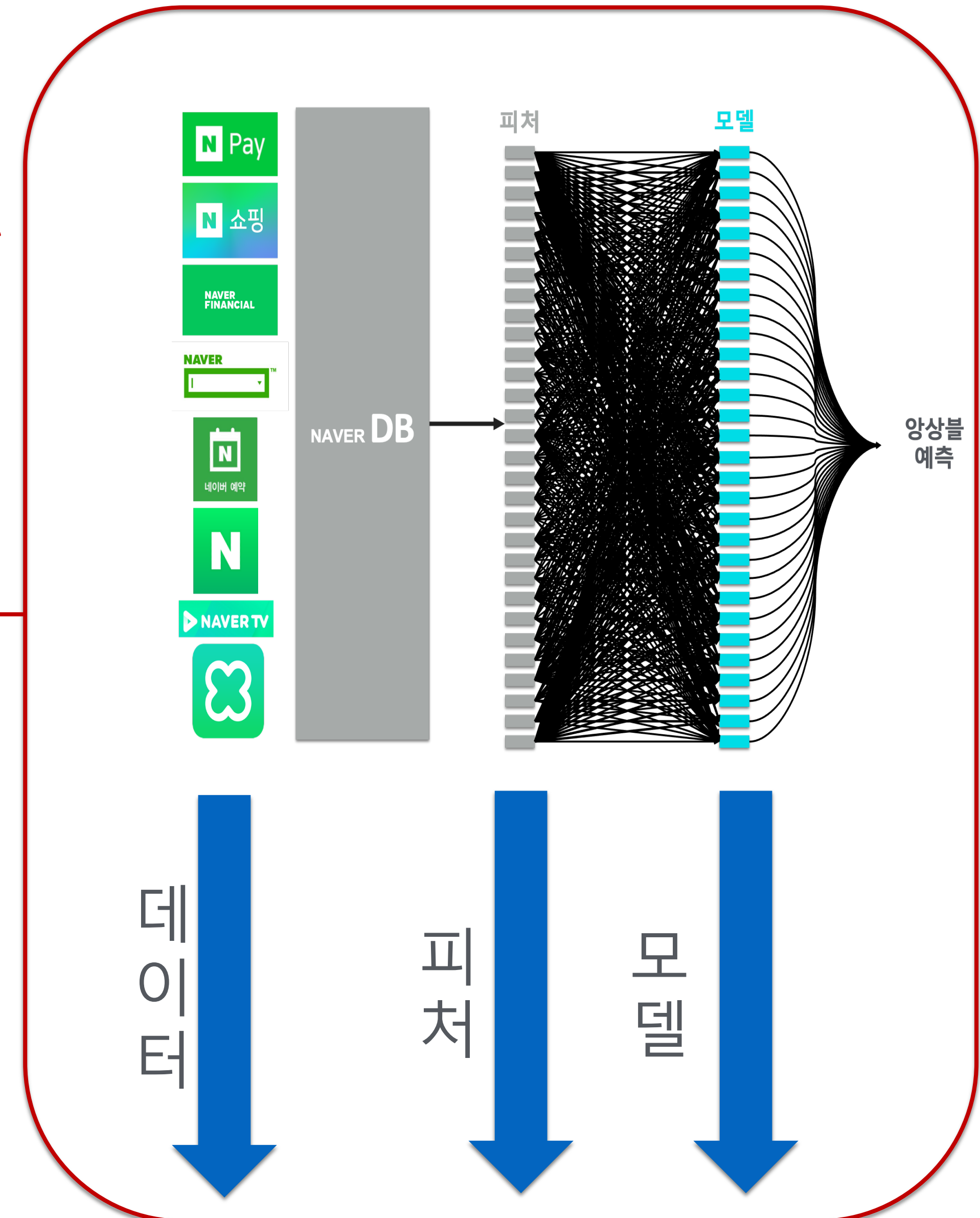
- HyperCLOVA의 효율성: SOTA모델 개발에 4-5개월의 리서치가 필요했지만 Ensemble SOTA는 HyperCLOVA ensgen 릴리즈 이후 **4일** 정도 만에 도달할 수 있었음.

- 문제의 전환: SOTA찾기 -> 다양한 모델 만들기  
(computing power!)

# 4.8 HyperCLOVA의 미래: computing power로 비즈니스 문제 해결



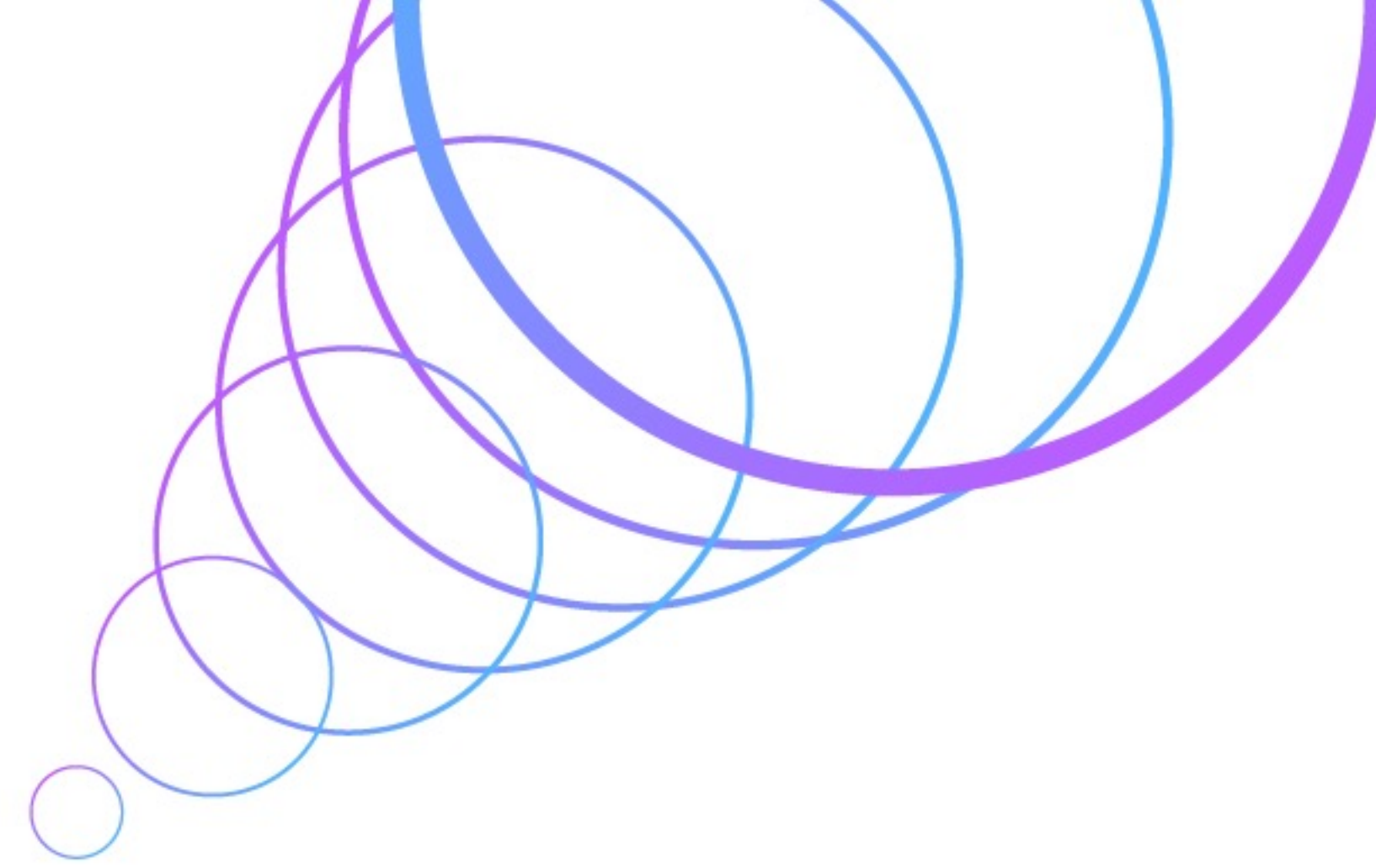
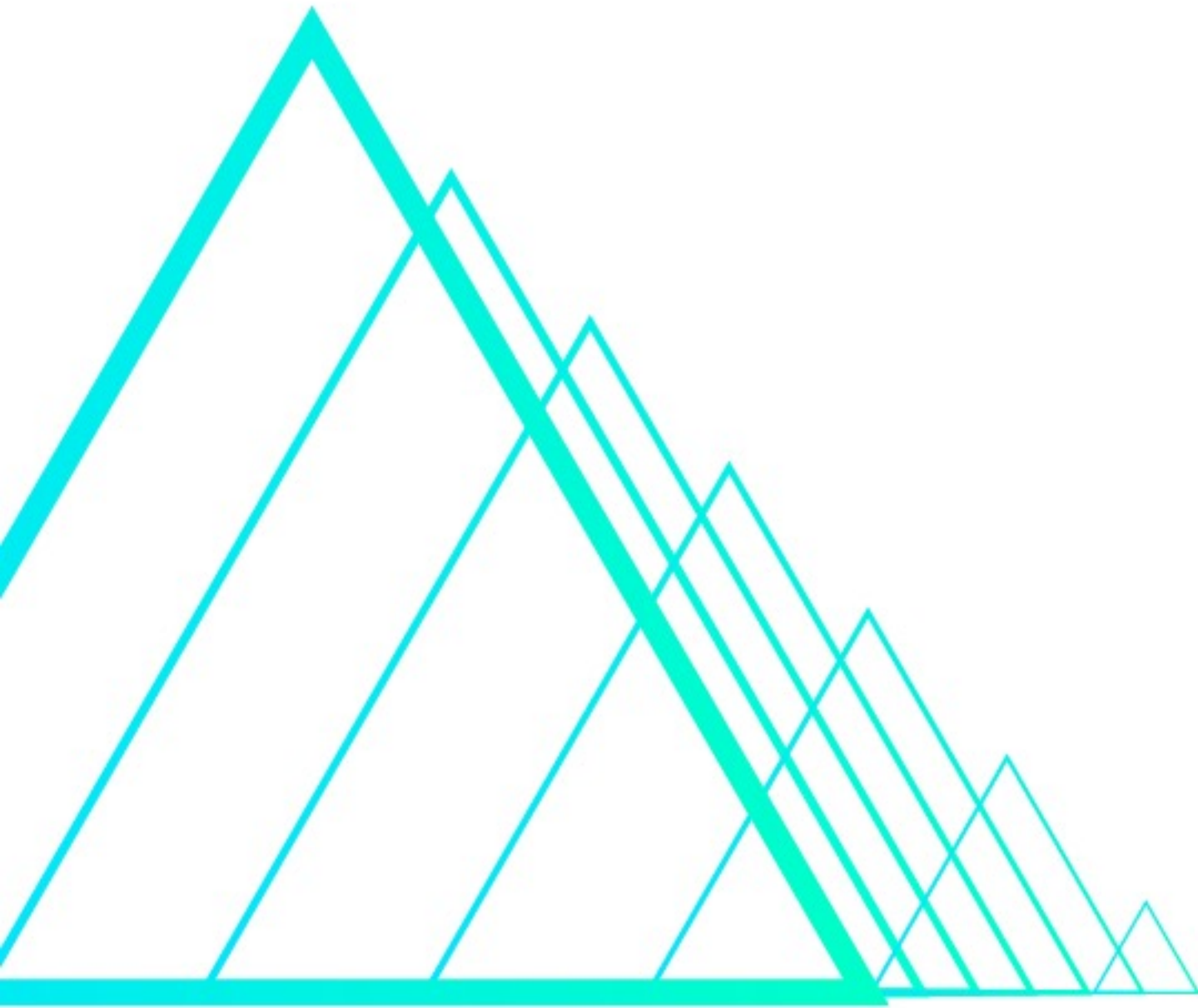
## HyperCLOVA Framework





We are hiring..

**clova-jobs@navercorp.com**



Thank You

