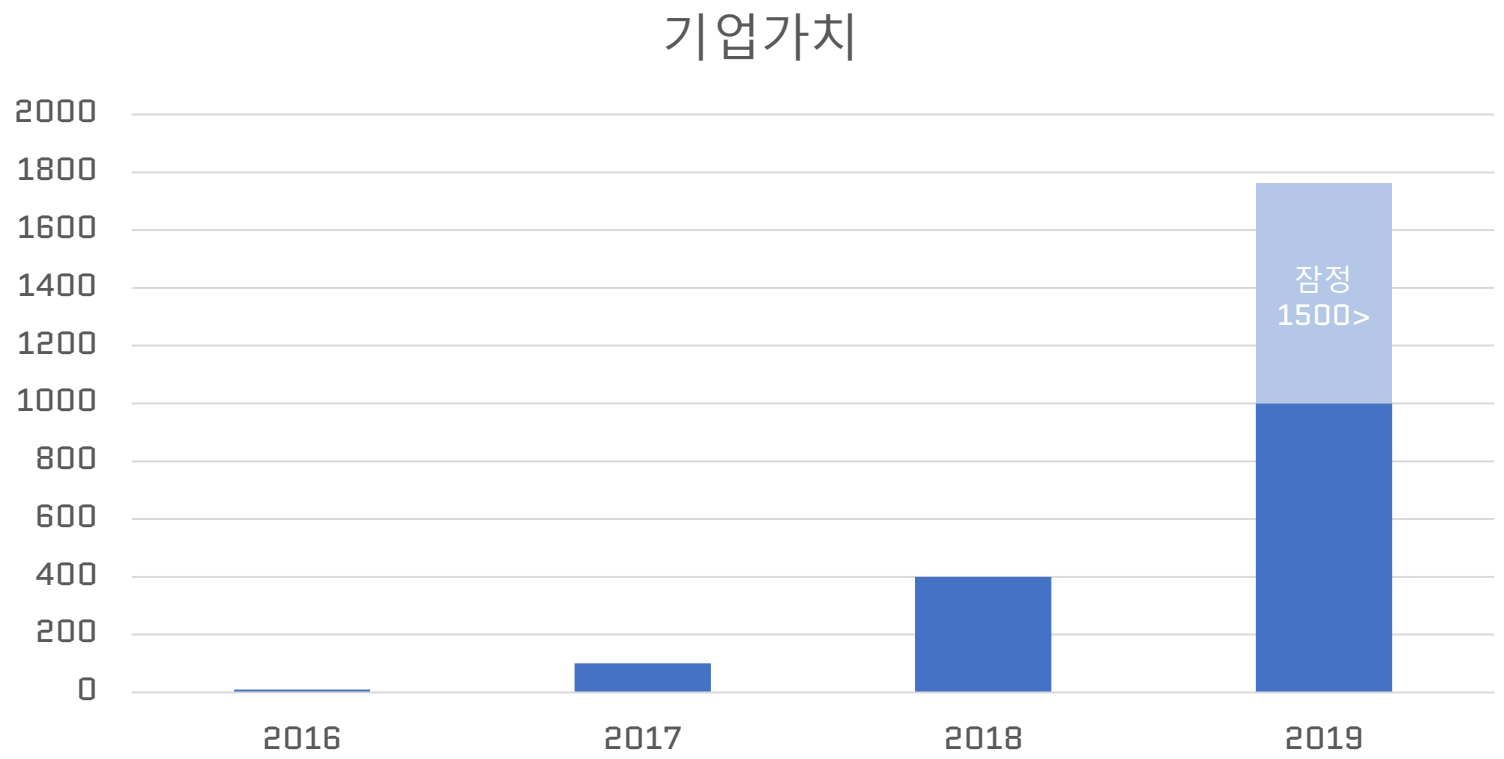
Investment & Deep Learning

QRAFT TECHNOLOGIES, INC.

March 2019

STRICTLY CONFIDENTIAL

크래프트테크놀로지스는 빠르게 성장하는 금융 AI 회사입니다



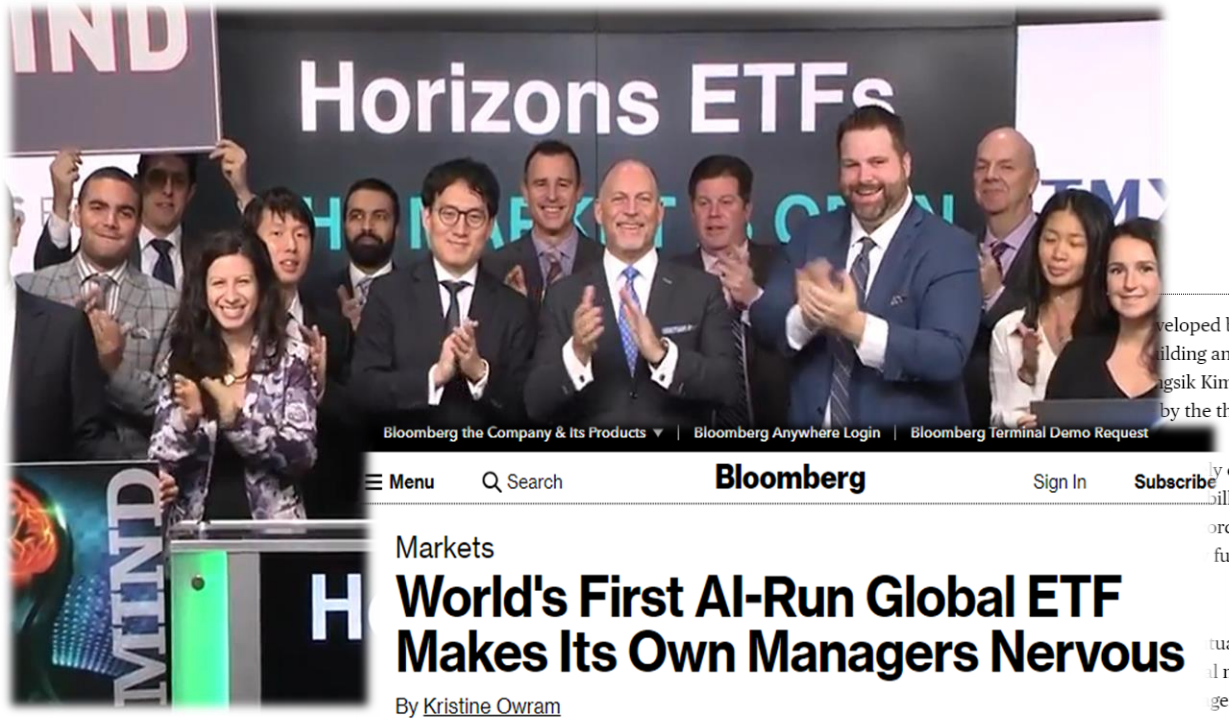
Market Share

국내 최대의 로보어드바이저 서비스 공급사

로보어드바이저 인더스트리 총 자산

크래프트 엔진 적용

세계 최초 글로벌 AI ETF 상장 & 국내 최초 AI ETF 뉴욕증시 상장 예정



기타

韓 금융AI 스타트업, AI ETF 뉴욕증권거래소 상장앞둬

최종수정 2019.02.11 09:56 기사입력 2019.02.11 09:18 댓글 쓰기

크래프트테크놀로지스 "ETF 2종 빠르면 75일내 NYSE 상장"

- "로또 찾지마!" 98% 모르는 3가지..
- 오빠 너무 단단해! 대체 뭘 먹은거야

[아시아경제 문채석 기자]금융 AI 스타트업 크래프트테크놀로지스(크래프트)가 인공지능(AI)으로 운용하는 상장지수펀드(ETF) 2종이 뉴욕증권거래소(NYSE) 상장을 앞두고 있다고 11일 밝혔다.

크래프트 측에 따르면 AI 기반으로 운용되는 ETF가 미국 거래소에 상장되는 것은 국내 최초다.

Bloomber

veloped by South Korea
ilding an office in Toror
gsik Kim said in an em
by the third quarter, he

y competitive an
illion ($21 billion
rding to Nationa
funds launched l

tual funds two to
l mutual fund pr
gement Ltd., Mac

Hawkins also expects growing competition from the big banks that could eventually result in "a big shift in market share," but said the Canadian ETF space is growing fast enough to boost most players in the near term. Horizons plans to launch about 10 new funds this year.

"We think we're in the heel of a hockey stick, we haven't hit the handle yet," he said, referring to shape. "We're still going to grow and we're still going to grow nicely."

UBS Is Curbing Some C
After Banker Detained

BUSINESSWEEK
Biohackers Are Implan
From Magnets to Sex T

AXE Challenge

국내 최초 AI 주문집행 시스템 개발

- 실거래 성과 측정 기간: 2018.11.14 ~ 22 (총 7거래일)
- 총매수금액: 2,536,275,830 원 (코스피 200 중 70종목)
- 일평균 매수금액: 362,325,119 원
- 시장 VWAP 대비 총 절감금액 : 1,217,412 원
- 대회기간 평균 절감 비율: **VWAP 대비 5bps 절감**

AI ETF

NVIDIA GTC 2018

문효준

現) 크래프트 테크놀로지스 AI리서치팀 팀장

1. 강화학습 기반 주문집행알고리즘 AXE 프로젝트 PM
2. US AI Enhanced ETF PM (4월초 뉴욕증시 상장)
3. Deep Asset Allocation PM (연기금 프로젝트)
4. AXE for Scheduling (Currency, Commodity) PM
5. 로보어드바이저 고도화 프로젝트
6. NLP 기반 US 인덱스 프로젝트

금융 + 딥러닝 = ?

왜 대부분은 기업에서 금융과 딥러닝을 접목시키는데 실패할까요?

Garbage In, Garbage Out

금융 데이터는 딥러닝을 적용하기에 최악의 구조

문제점 1. 시계열 Feature 자체의 노이즈

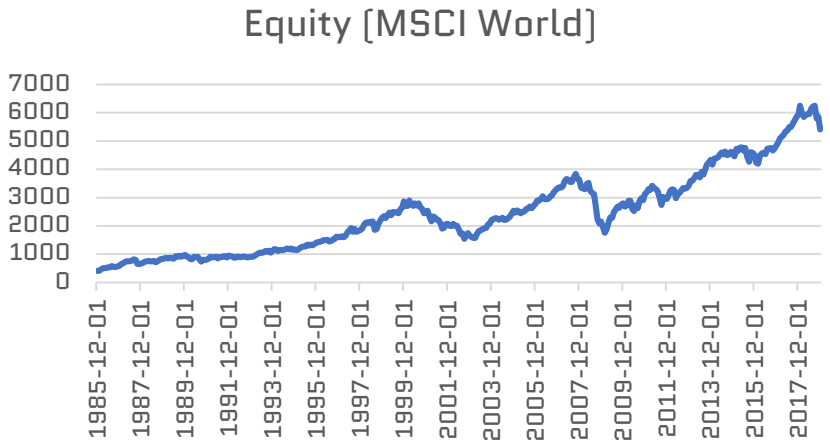
예시) 수익률 데이터

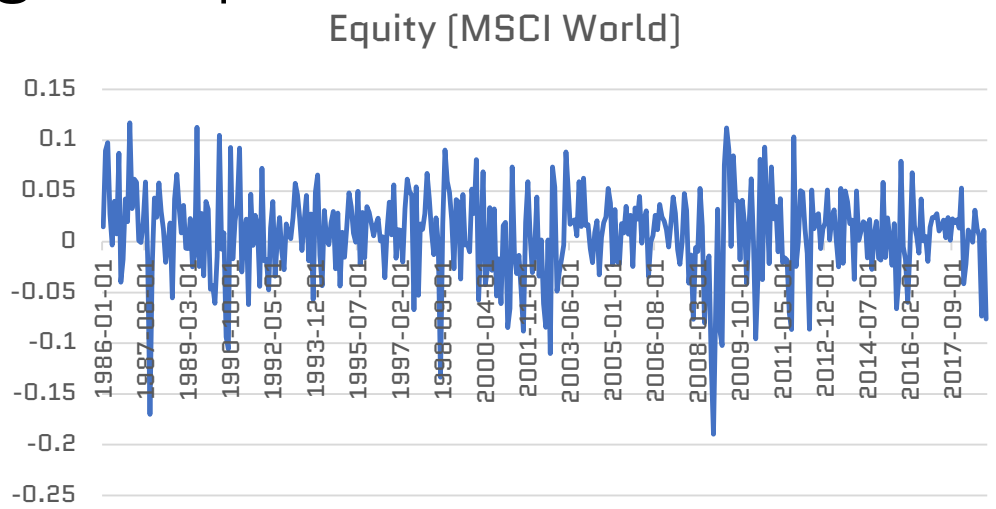
일반적으로 주가 모델링은 기하브라운모형 (GBM)을 가정하지만,
근본적으로 AR(1)과 크게 다르지 않으므로 편의를 위해 AR(1) 모형에서 시작해 조금만 변형해봅시다.

$$y_{t+1} = y_t + x_t + \epsilon_t$$

딥러닝 모형이 포착해야 함

다음 주가 = 현재주가 + 정보 + 노이즈





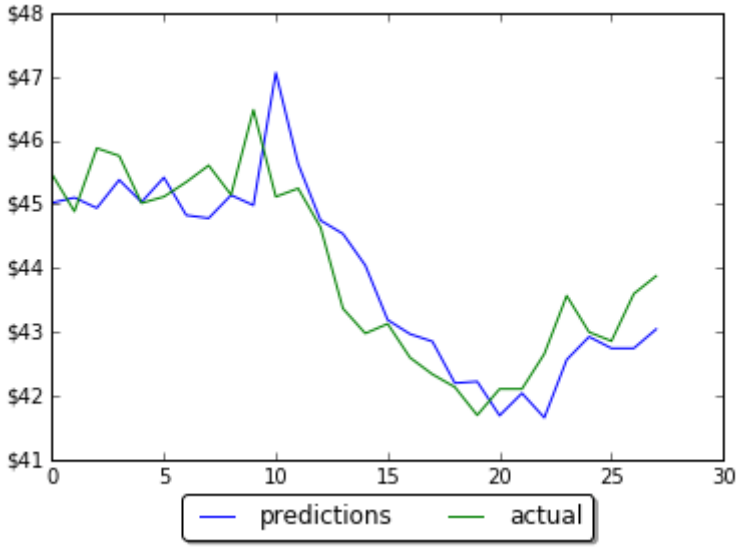
하지만 ϵ_t가 차지하는 요소 > x_t가 차지하는 요소
노이즈 〉 정보량

예시) 수익률 데이터

$$y_{t+1} = y_t + x_t + \epsilon_t$$

다음 주가 = 현재주가 + 정보 + **노이즈**

노이즈로 인해 상대적으로 $y_{t+1} = y_t + \epsilon_t$ 처럼 되고 결과적으로 다음값에 대한 최선의 예측값은 현재값.



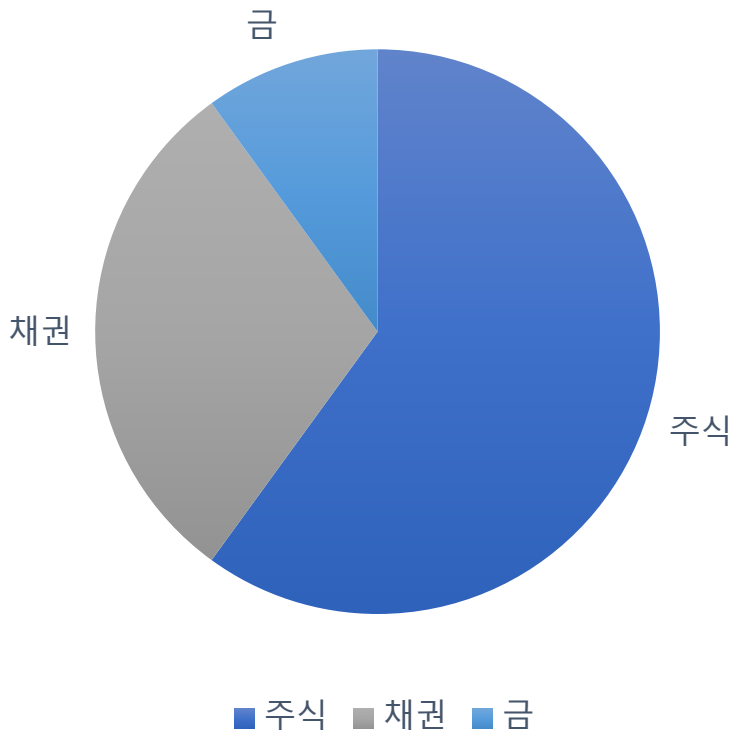
LSTM을 이용한 주가 예측
잘 예측하는 것 같지만,
실질적으로 오른쪽으로 Lagging

문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

예시) 자산배분문제

Q : 요즘 같은 경제 상황에서는 어떤 자산군에 어떻게 투자 해야할까요 ?

A : 음… 주식 60%, 채권 30%, 금 10% ?



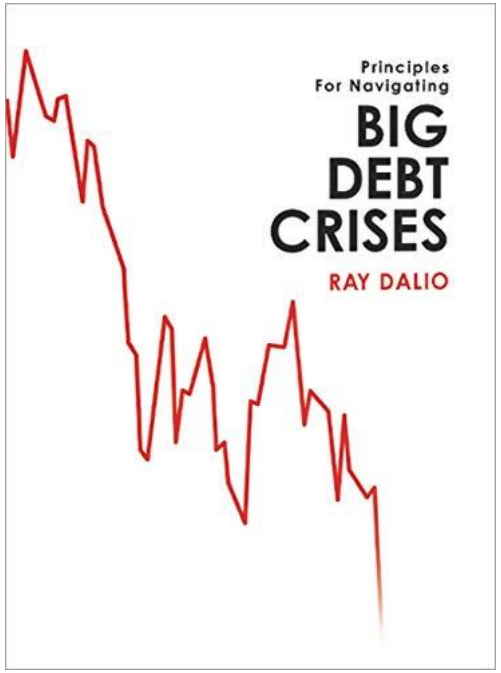
자산배분을 위해 고려해야할 요소

금, 채권, 주식, 리츠, 원자재, … 등 수많은 자산군 데이터
금리, 인플레이션, 장단기금리차, … 등 수많은 매크로 데이터

필요한 High Level Feature ?

자산군 모멘텀 효과
자산군 평균회귀 효과
확장적 통화정책, 긴축적 통화정책 분류
단기부채사이클, 장기부채사이클 파악
…

→ 주로 Monthly Frequency 데이터에서 추출

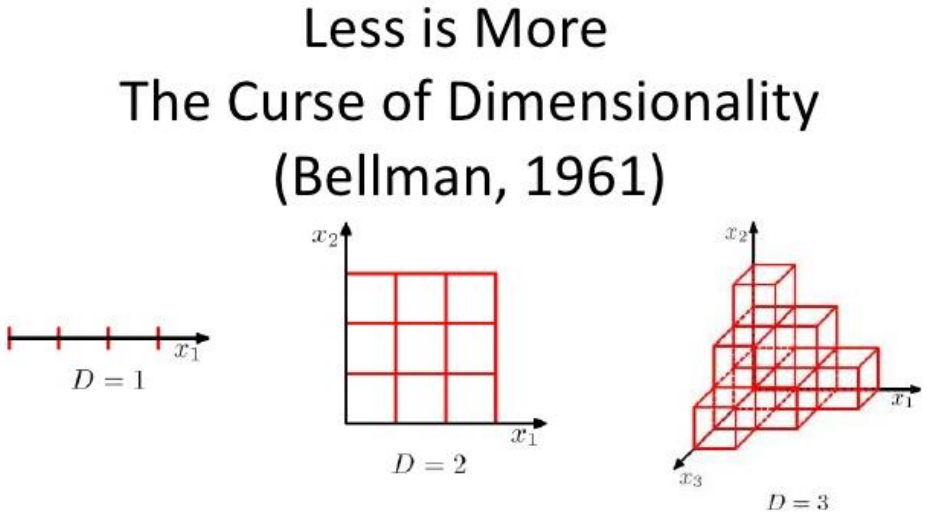


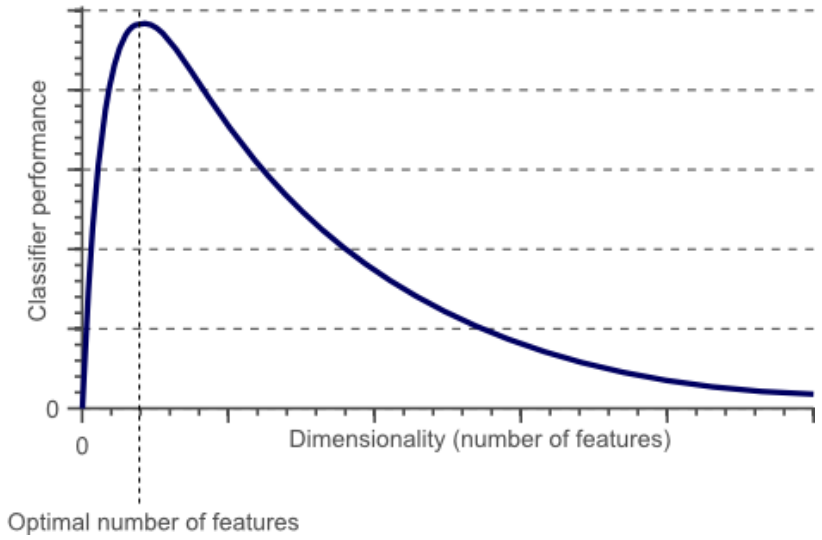
40년 데이터 = 겨우 480개의 Sequence 길이

(Monthly Frequency 기준)

고려할 수 있는 요소는 수십 ~ 수백개인데, 고려할 수 있는 데이터 길이는 너무 짧음

→ 차원의 저주





문제점 3. 문제점 1과 문제점 2로 인한 Overfitting

문제점 1. 시계열 Feature 자체의 노이즈

문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

문제점 3. 문제점 1과 문제점 2로 인한 Overfitting

어떻게 해결할 것인가?

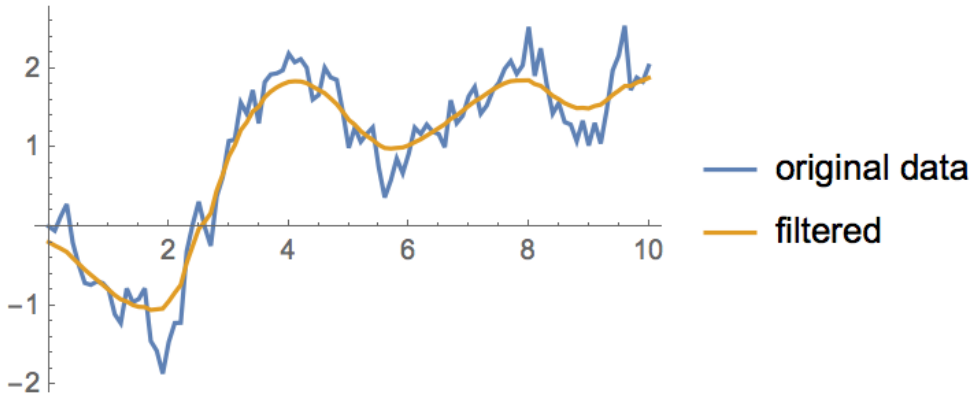
차례대로 해결해봅시다.

문제점 1. 시계열 Feature 자체의 노이즈

Time-series denoising

1. Moving Average (MA, EMA, …)

2. Bilateral Filter



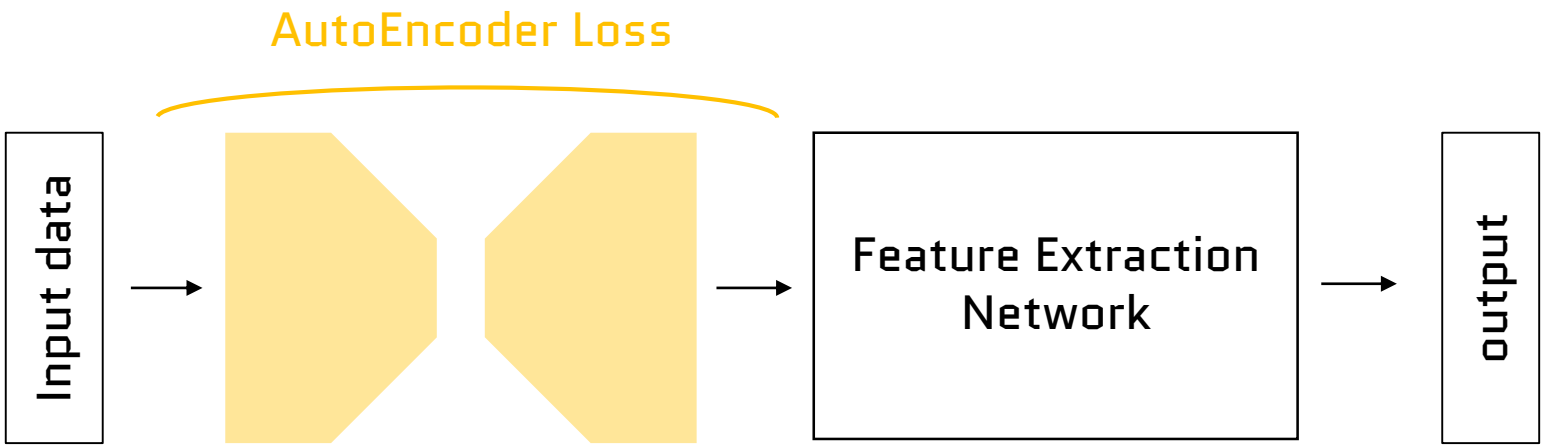
문제점 1. 시계열 Feature 자체의 노이즈

Time-series denoising

Moving Average, Bilateral Filter → 학습 개선 효과 존재

하지만 학습 과정에서 자동적으로 노이즈를 제거할 수 있는 방법은 없을까?

CNN Stacked AutoEncoder 기반 Denoising Module

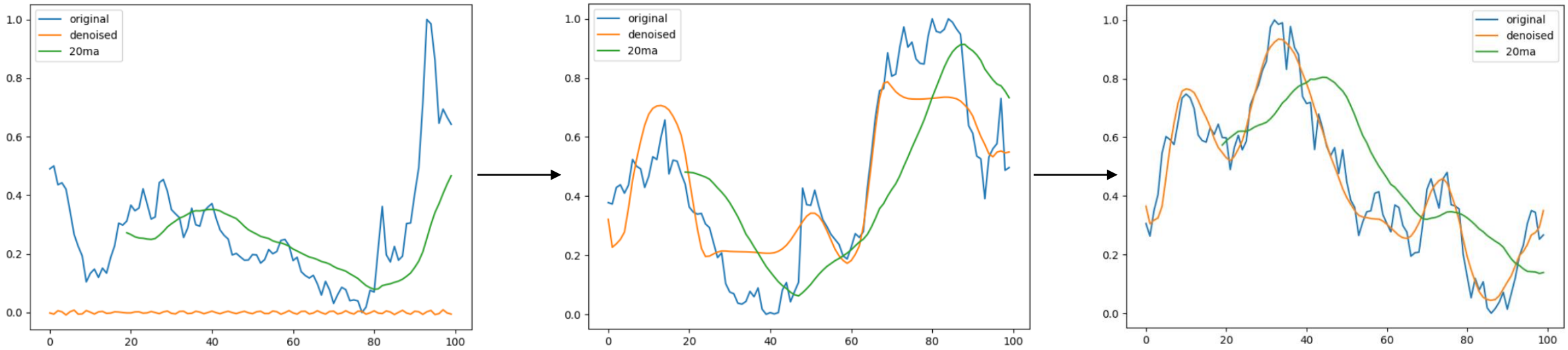


기존 지도학습에서의 Loss + AutoEncoder Loss → Gradient Descent

문제점 1. 시계열 Feature 자체의 노이즈

Time-series denoising

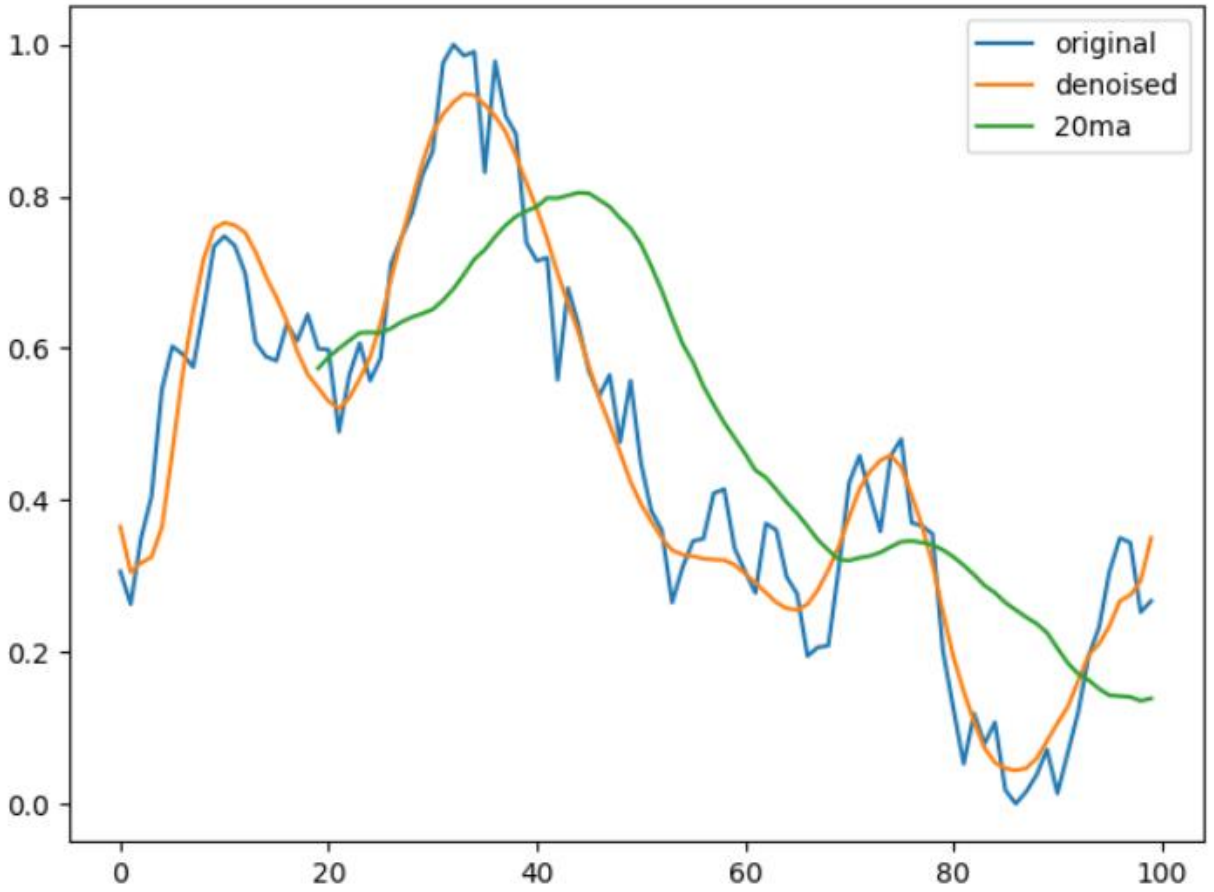
CNN Stacked AutoEncoder 시각화



문제점 1. 시계열 Feature 자체의 노이즈

Time-series denoising

CNN Stacked AutoEncoder 시각화



문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

데이터 생성 ?

→ GAN 기반 데이터 생성. 개별적인 생성은 가능하지만, 전체 시계열의 상관성을 고려한 생성은 어려움

데이터 생성도 어렵고, 실제 데이터가 부족한데 어떻게 모델을 학습시킬 것인가 … ?

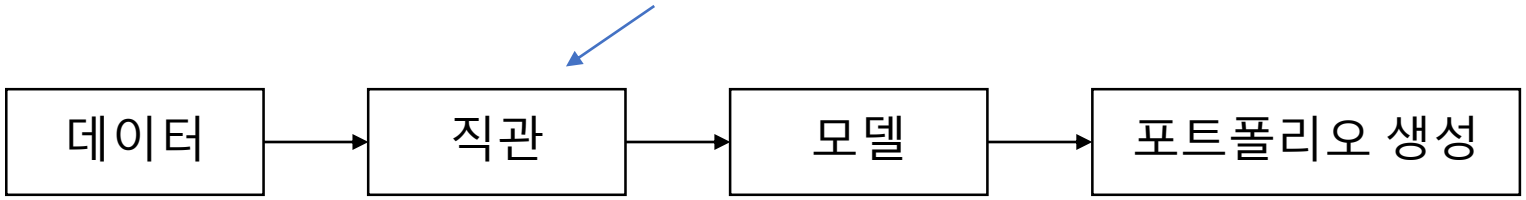
거꾸로 생각해보기

이런 상황에서 기존 퀀트들은 모델을 어떤 식을 만들었는가?

문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

기존 퀀트가 모델을 만드는 방식

(간접적으로라도) 경제적 함의점을 내포하는 모델 설계

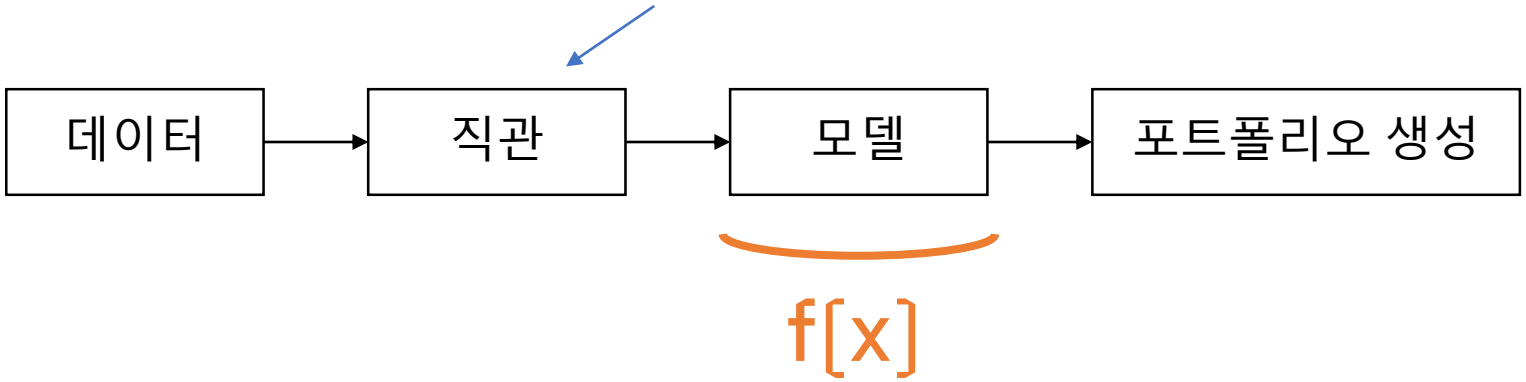


경제적 함의점까지 고려하는 직관 자체를 모델링 하는 건 (아마도) 불가능

문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

기존 퀀트가 모델을 만드는 방식

(간접적으로라도) 경제적 함의점을 내포하는 모델 설계



f[x]는 주로 선형적인 모델

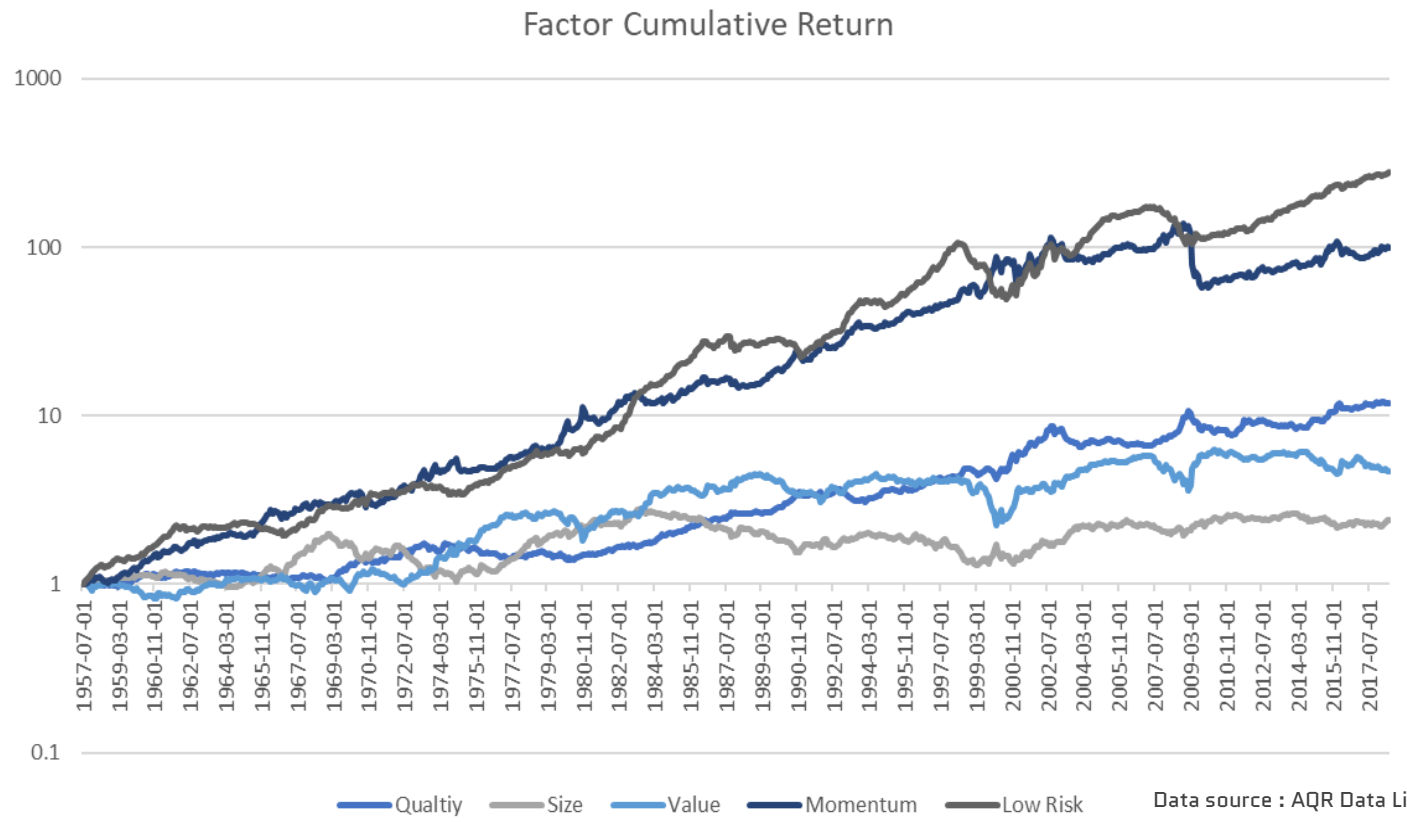
예를 들어 Momentum Factor 모델링

12 Month Price Data → Model → 12-1M → Momentum

단순 선형 모형에서 딥러닝으로 최적의 함수 추정

문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

사례) Dynamic Factor Rotation (Timing)



Data source : AQR Data Library

일정하지 않은 팩터 알파

문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

사례) Dynamic Factor Rotation (Timing)

No.	Thesis	Cause	Factor
1	Avramov, D., Cheng, S., & Hameed, A. (2016). Time-varying liquidity and momentum profits. Journal of Financial and Quantitative Analysis, 51(6), 1897-1923.	High Liquidity	Momentum ↑
2	Zhang, L. (2005). The value premium. The Journal of Finance, 60(1), 67-103.	High / Low Economy Activity	Value ↑↓
3	Jensen, G. R., & Mercer, J. M. (2002). Monetary Policy and the Cross-Section of Expected Stock Returns. Journal of Financial Research, 25(1), 125-139.	Monetary Expansion / Contraction	Value ↑↓
4	Black, A. J., Mao, B., & McMillan, D. G. (2009). The value premium and economic activity: Long-run evidence from the United States. Journal of Asset Management, 10(5), 305-317.	Economic Expansion / Contraction	(Value – Growth) ↑↓
		Money Supply Increase / Decrease	Value ↑↓
		Interest Rate Increase / Decrease	(Value – Growth) ↑↓
5	Asness, C. S., Frazzini, A., & Pedersen, L. H. (2017). Quality minus junk.	Recession, Crises	(Quality – Other Factors) ↑
6	Barroso, P., & Santa-Clara, P. (2015). Momentum has its m oments. Journal of Financial Economics, 116(1), 111-120.	High Volatility	Momentum ↓

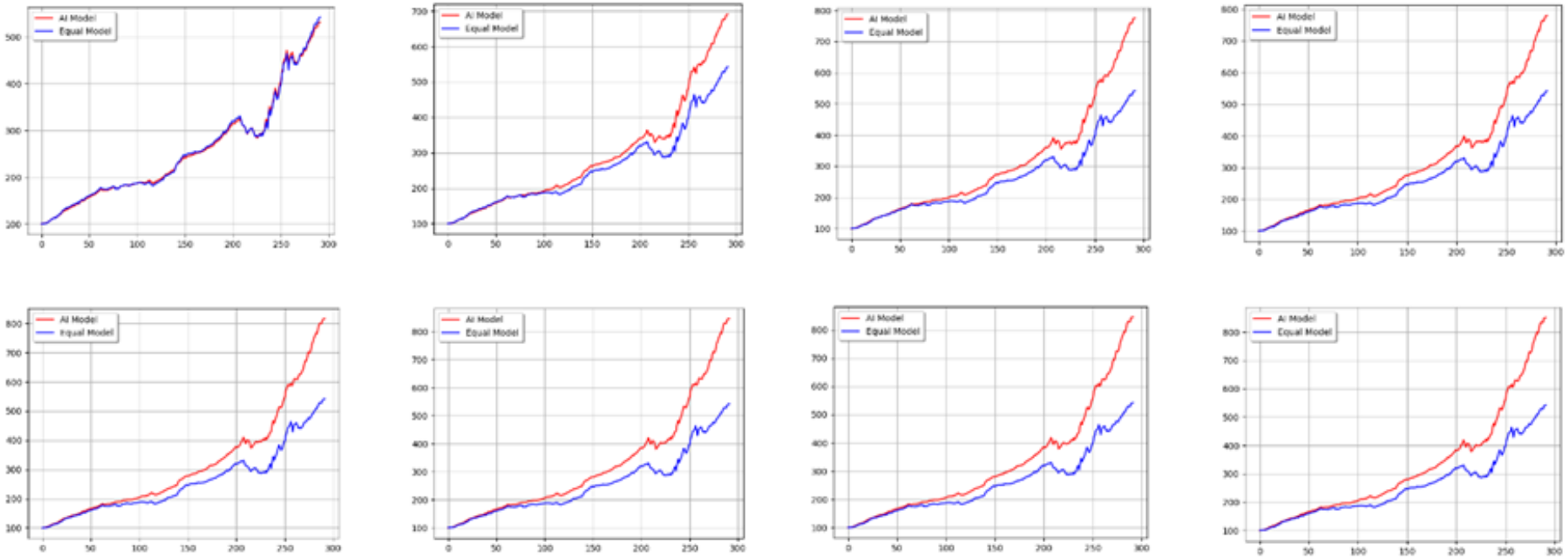
문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

사례) Dynamic Factor Rotation (Timing)

Simulation with train data set during training
(1980 ~ 2005)

- : AI Factor Rotation Model
- : Equal Factor Model

At first, it is similar to equal factor weight model



As time goes by, AI model learns to optimally allocate factors with train data set

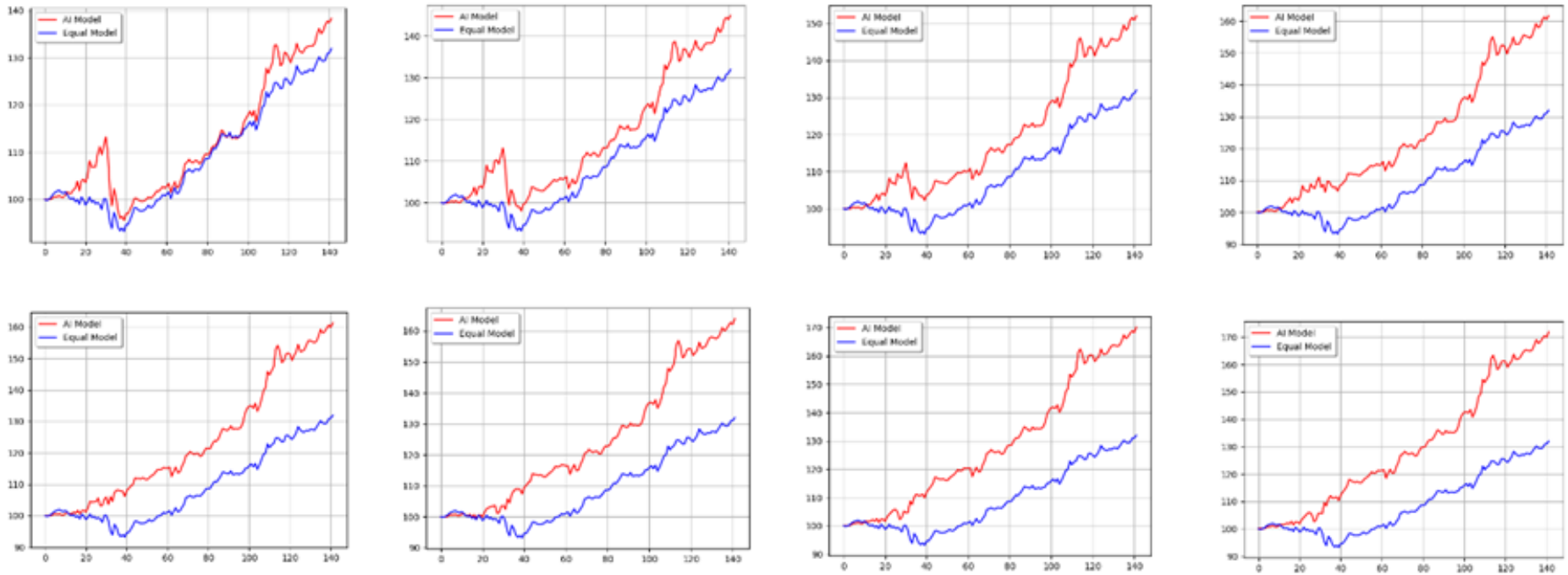
문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

사례) Dynamic Factor Rotation (Timing)

Simulation with test data set during training
[2006 ~ 2018]

- : AI Factor Rotation Model
- : Equal Factor Model

At first, it is similar to equal factor weight model

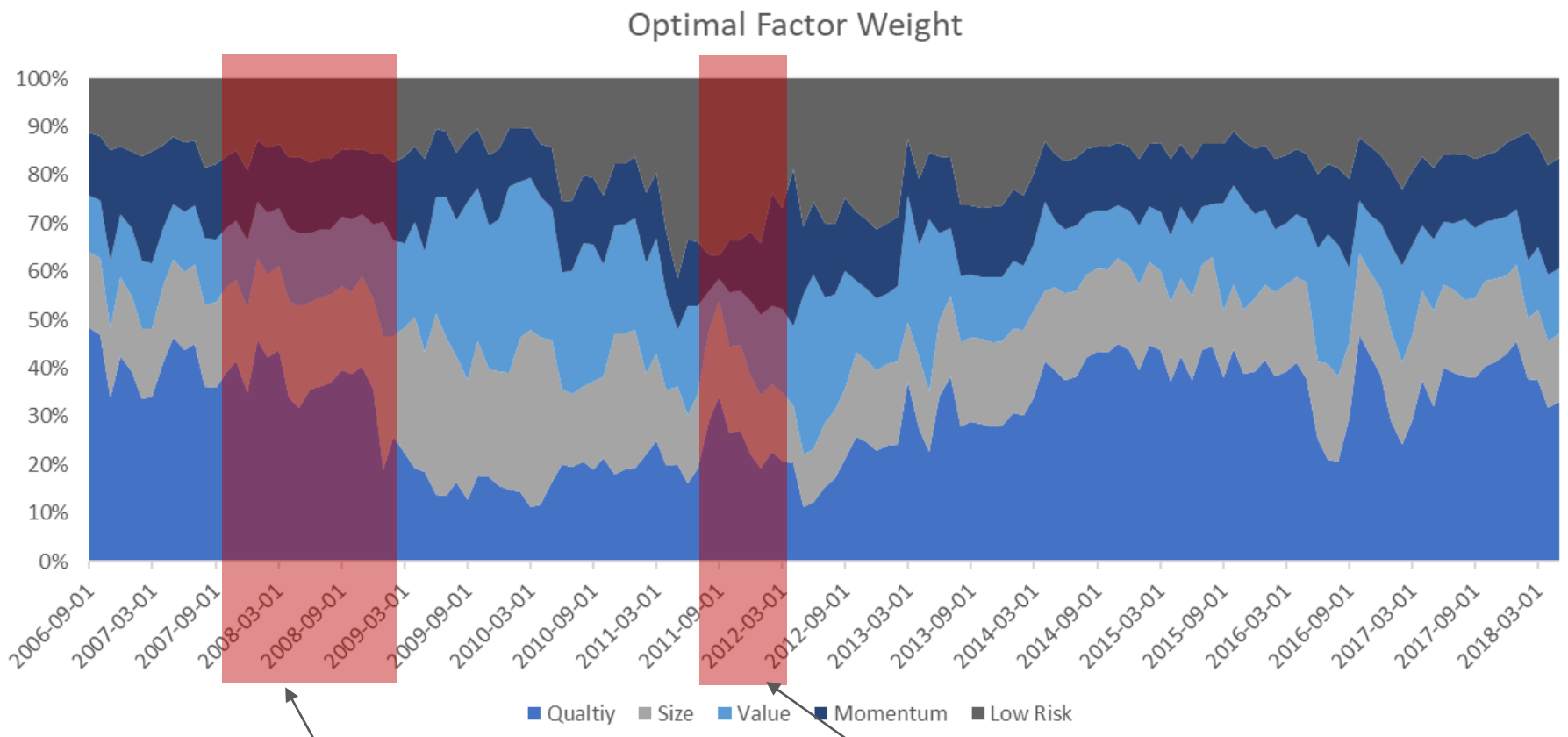


As times goes by, AI model learns to optimally allocate factors with test data set

Paired t-test p-value : 0.00263132

문제점 2. 시계열 Feature 종류 대비 짧은 시계열 길이

사례) Dynamic Factor Rotation (Timing)



August 2011 (Black Monday) : **High '<u>Low Risk</u>' weight**

2007 – 2008 (Financial Crisis) : High '**<u>Quality Factor</u>**' weight

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

문제점 1과 문제점 2를 어느 정도 해결하더라도 여전히 Overfitting 발생

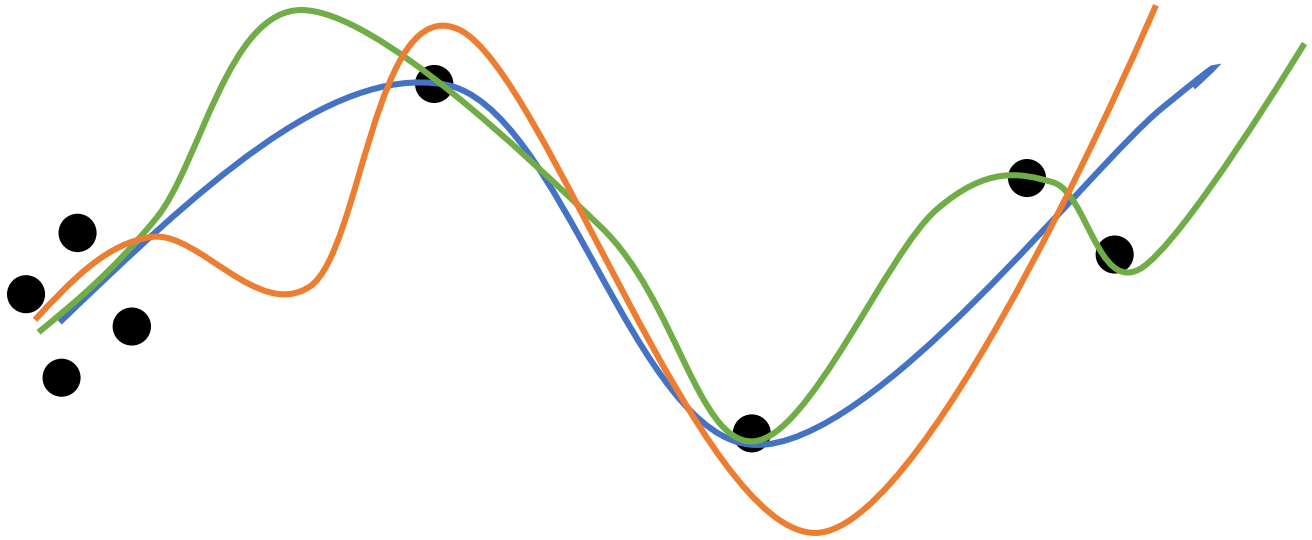
그렇다면 어떻게 Overfitting을 최소화할 것인가?

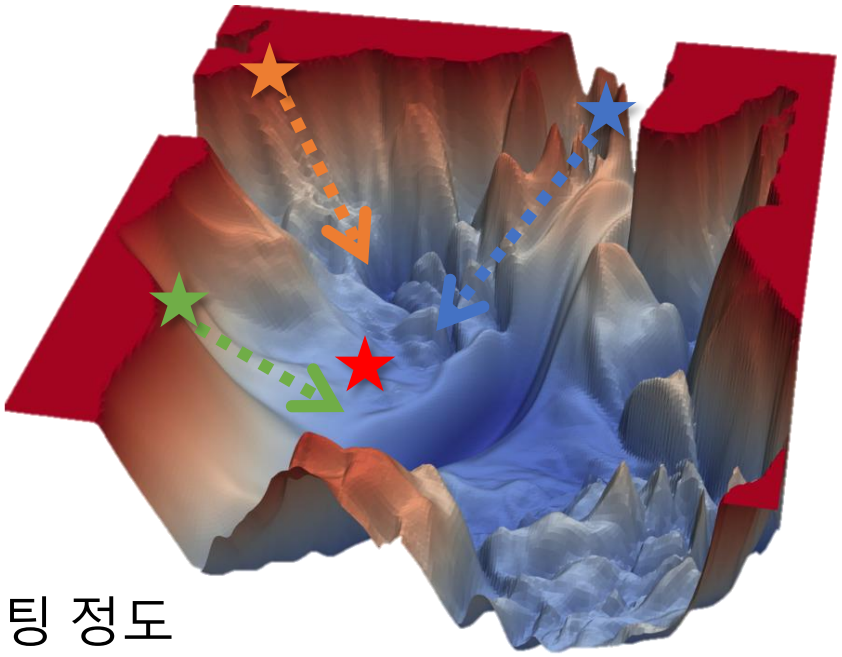
1. Asynchronous Multi Network Learning

2. Bayesian Inference

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

1. Asynchronous Multi Network Learning



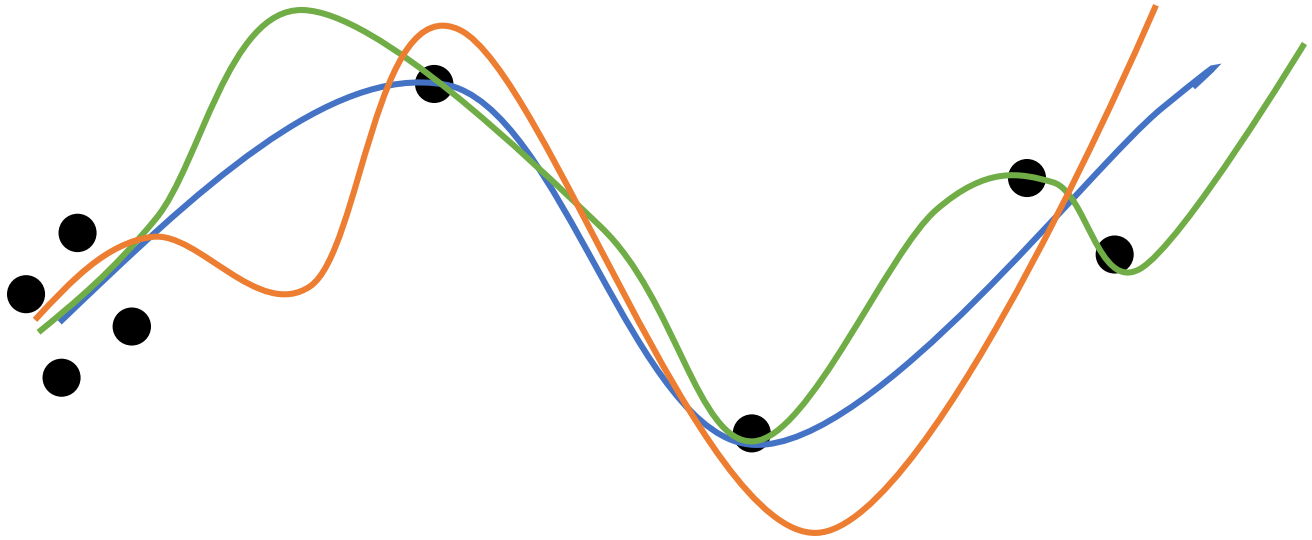


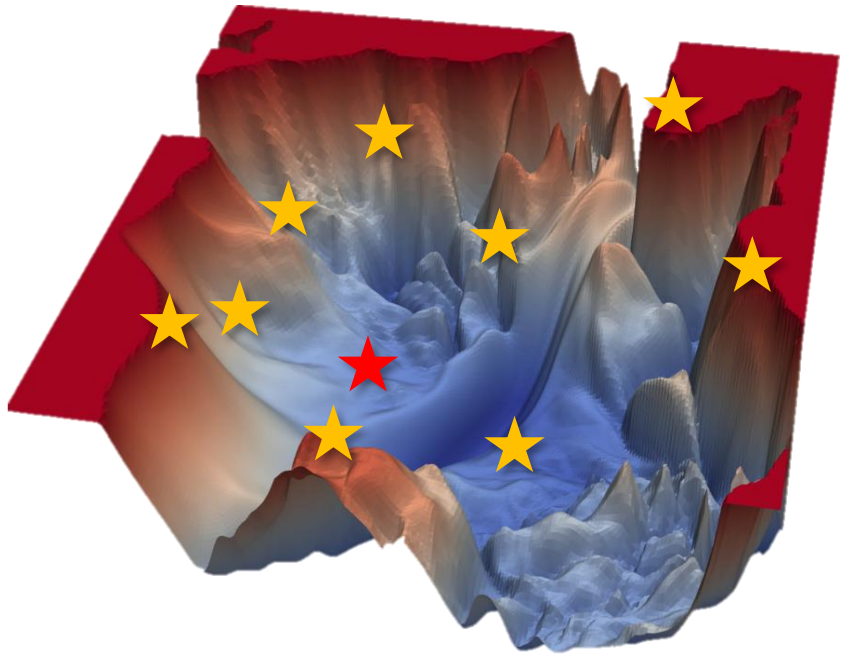
비슷한 **Cost**이지만 다른 오버피팅 정도
→ **Weight Initialization**에 따라 다르게 수렴

일반적으로 데이터가 부족해 단순 **L1, L2 Norm**을 적용시킬 수 없음

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

1. Asynchronous Multi Network Learning





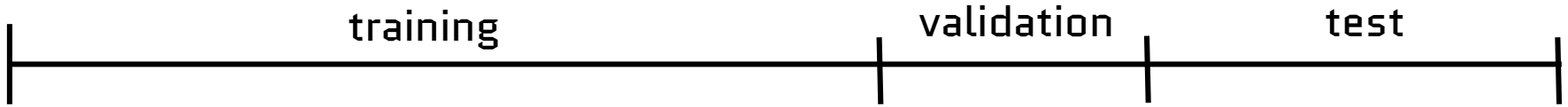
Thread가 개별적으로 초기화된 각 네트워크를 가지고, 여러 네트워크를 **동시에** 학습 시작

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

1. Asynchronous Multi Network Learning

Training 중 Test data를 가지고 오버피팅 체크를 하게 되면 Look-ahead bias 발생

따라서 별도의 validation data로 오버피팅 감지

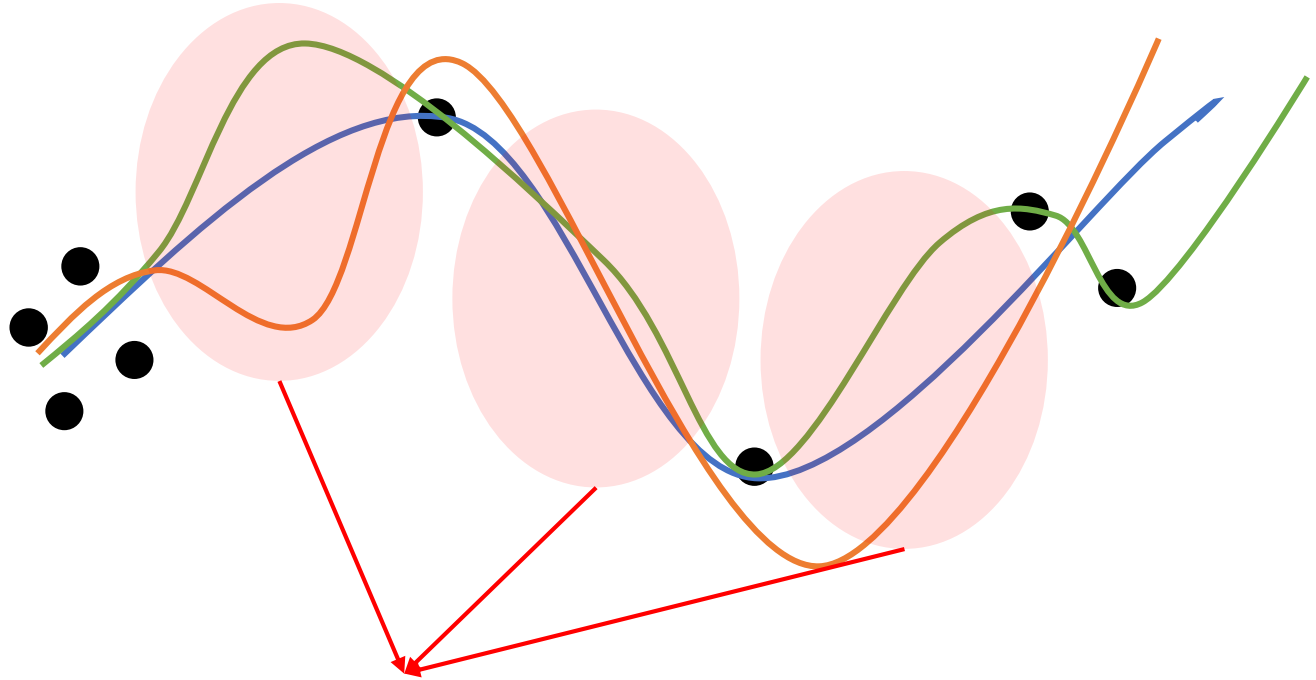


1. N개의 쓰레드로 네트워크 개별 초기화
2. K iteration 후 validation data로 오버피팅 검증 후 하위 X개 제거
3. 새롭게 초기화된 X개의 네트워크 투입
4. Validation에서 원하는 수준까지 수렴할 때까지 반복
5. 잔여모델 Test에서 앙상블 후 모델 최종 검증

→ 효과적인 오버피팅 방지 [전 Factor Rotation 사례에서 검증]

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

2. Bayesian Inference



관찰하지 못한 데이터구간 → Validation에서도 관찰이 불가능하다면,
오버피팅보다는 **모른다고 결과를 내는 것이 효과적**

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

2. Bayesian Inference

어떻게 모른다고 결과를 낼 것인가? Uncertainty Quantification

1. Monte Carlo Dropout

(Gal, Y., & Ghahramani, Z. (2016, June). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In international conference on machine learning (pp. 1050-1059)

2. Monte Carlo Batch Normalization

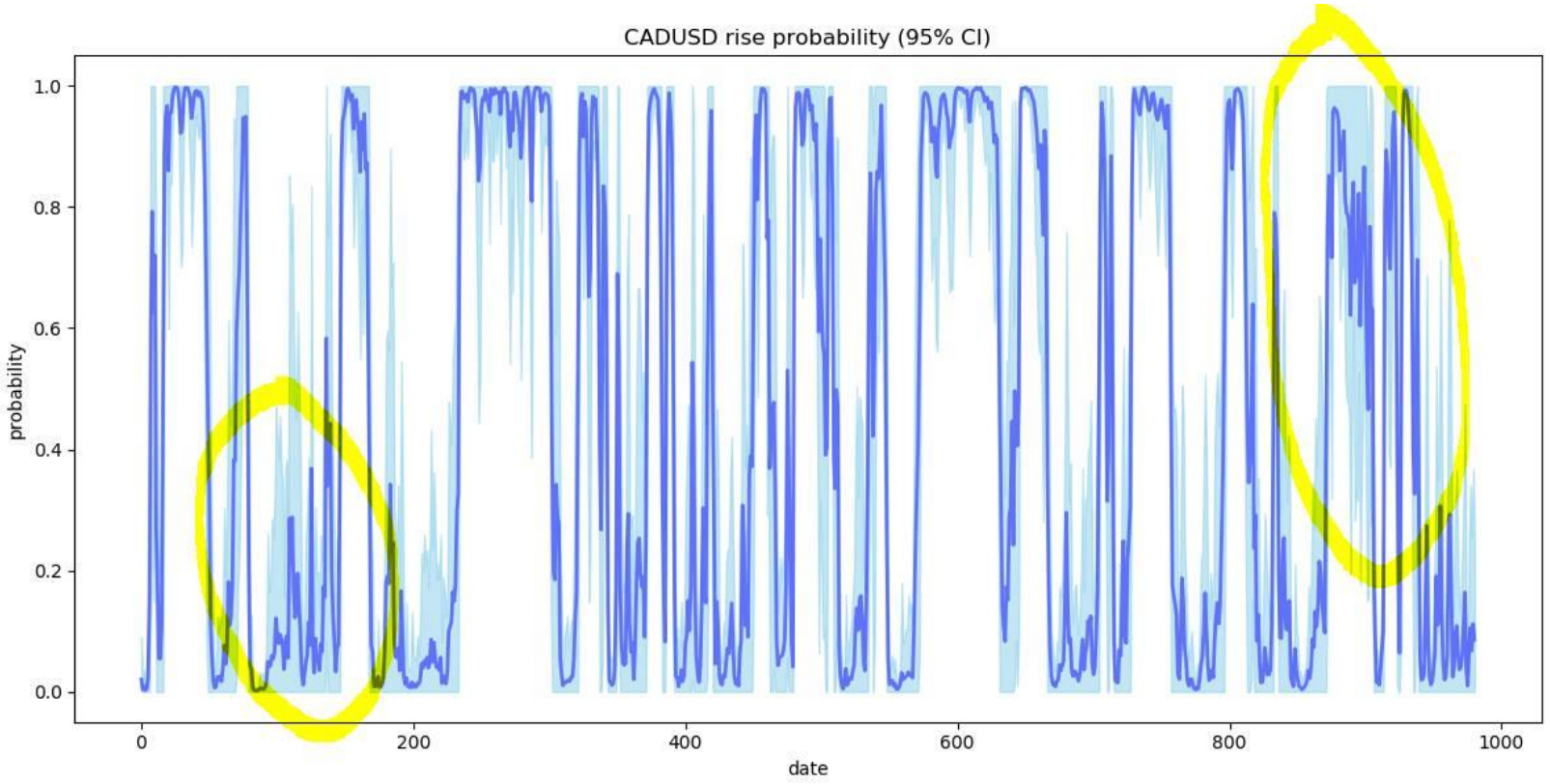
(Teye, M., Azizpour, H., & Smith, K. (2018). Bayesian uncertainty estimation for batch normalized deep networks. *arXiv preprint arXiv:1802.06455)*

2. Deep Learning Regression + Gaussian Process Regression

선지도학습 후 GPR 학습 → 가장 심플하고 적용하기 간단

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

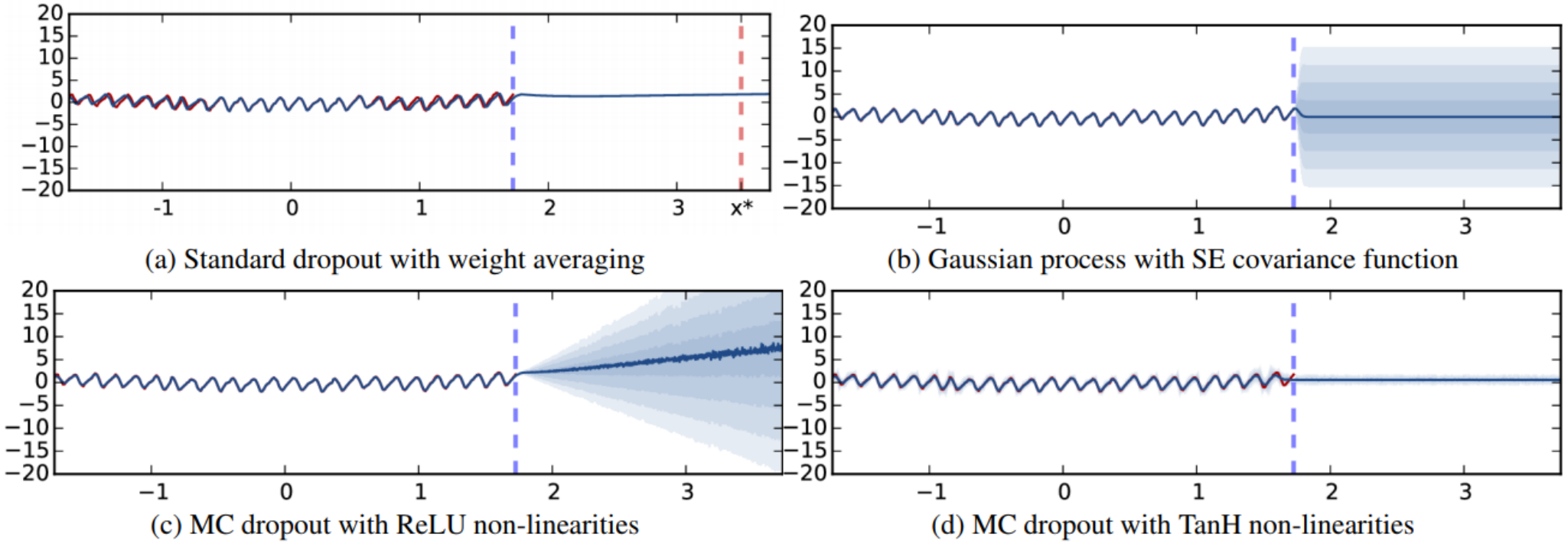
2. Bayesian Inference



Qraft Currency Hedge Engine – MC Dropout 적용 예시

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

2. Bayesian Inference



(a) Standard dropout with weight averaging

(b) Gaussian process with SE covariance function

(c) MC dropout with ReLU non-linearities

(d) MC dropout with TanH non-linearities

MC Dropout → Tau, Dropout rate, Activation에 따른 영향도 높다는 단점

문제점 3. 그럼에도 불구하고 발생하는 Overfitting

2. Bayesian Inference

어떻게 모른다고 결과를 낼 것인가? Uncertainty Quantification

1. Monte Carlo Dropout

(Gal, Y., & Ghahramani, Z. (2016, June). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In international conference on machine learning (pp. 1050-1059)

2. Monte Carlo Batch Normalization

(Teye, M., Azizpour, H., & Smith, K. (2018). Bayesian uncertainty estimation for batch normalized deep networks. *arXiv preprint arXiv:1802.06455*)

2. Deep Learning Regression + Gaussian Process Regression

선지도학습 후 GPR 학습 → 가장 심플하고 적용하기 간단

선형 독립적으로 노드들이 Representation Learning이 잘 됐다는 가정 하에 효과적인 GPR 학습 가능

THANK YOU

QRAFT