

딥러닝을 활용한 투자 알고리즘

CUAI 딥러닝 세미나
이혜지

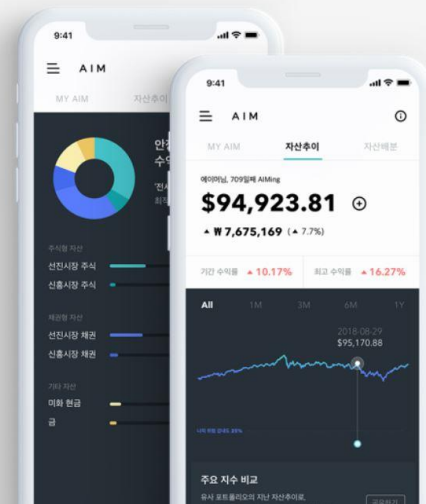
AIM 이란 무엇인가?



당신의 삶에 긍정적 가치를 더하다

상위 **1%** 자산관리
서비스

앱 다운로드



Q. AIM은 어떤 서비스인가요?



Q. AIM의 투자 알고리즘은 무엇을 기초로 만들어진 건가요?



AIM의 알고리즘인 '에스더'가 추구하는 투자전략은 펀드 매니저가 적극적인 종목 발굴과 매매를 통해 시장보다 높은 수익 실현을 추구하는 active 투자와 금융시장 전체(지수펀드)에 투자해 시장수익을 따라가는 passive 투자 양쪽의 장점을 모두 가지고 있습니다.

시장지수에 투자하는 ETF를 활용해 맞춤형 자산 포트폴리오를 구성하는 점은 비용 효율적인 passive 투자와 같지만, 시장국면의 변화를 감지하여 적절한 자산배분 변화를 주는 점은 active 투자의 속성에 가깝습니다.

단, 자산배분 변경 시 수익증대보다는 위험관리에 초점을 두어 '단기손실'에 민감한 개인투자자의 니즈에 최적화하였습니다.

즉, 고위험/고수익을 선호하는 투자자일지라도 시장사이클이 위험회피국면인 경우 포트폴리오 내 안전자산 비중을 높임으로써 위험을 낮추는 것은 물론, 안전자산으로의 투자수요 쏠림현상에 의한 초과수익까지도 기대할 수 있습니다.

석사학위논문

딥뉴럴네트워크를 이용한 중장기 투자전략
Long Term Stock Investment
Strategy with Deep Neural Networks

홍익대학교 대학원

컴퓨터공학과

유 다 빈

2018년 2월

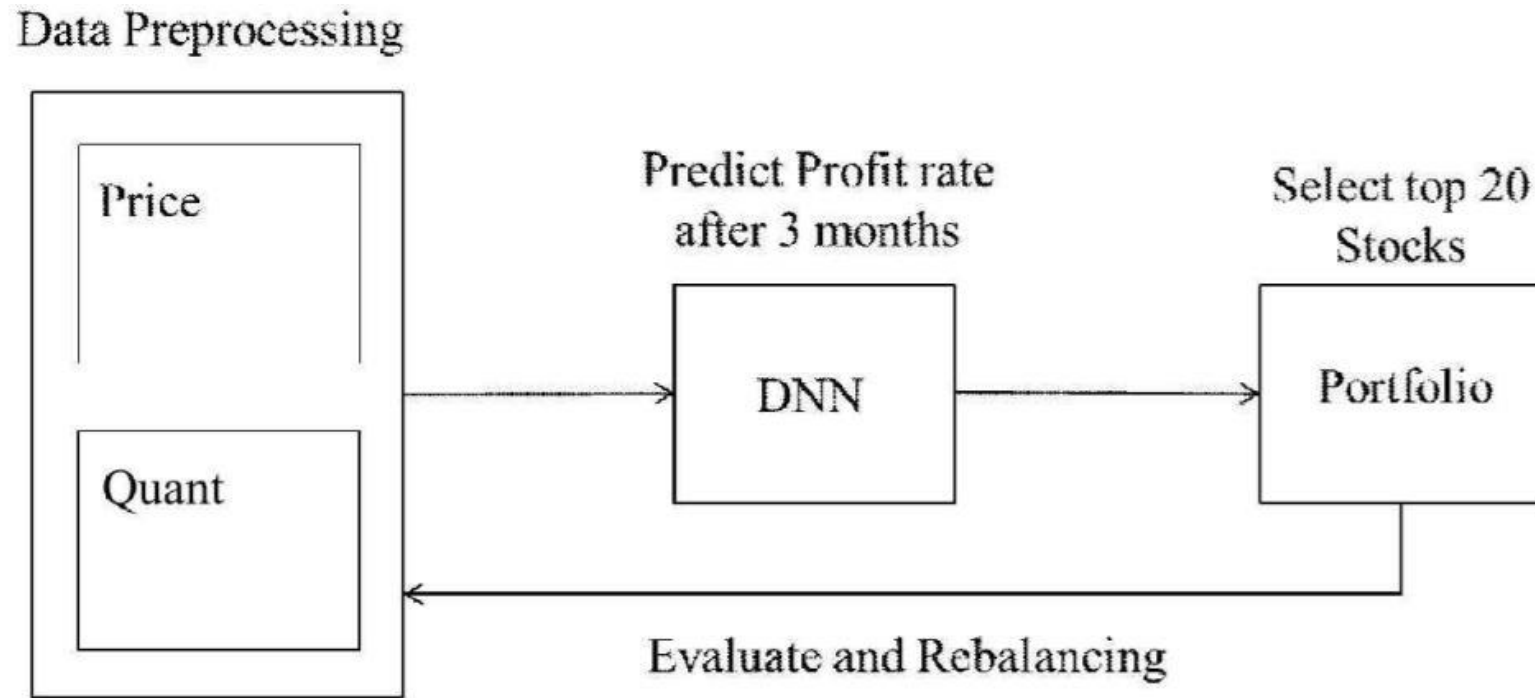


HONGIK UNIVERSITY

논문의 목표 및

- ✓ 딥러닝을 적용해 주식상승 예측
- ✓ 포트폴리오 기반의 중장기 투자 전략
 - 포트폴리오 전략은 중장기투자전략으로 성장할 기업들의 주식을 매입하고 주기적으로 종목을 변경해 수익 실현
- ✓ 3개월마다 종목 변경 -> 기업이 3개월마다 재무제표를 공개
- ✓ 투자 전략의 성공은 수익률과 MDD로 평가
 - Maximum Draw Down 는 최대 어느 정도의 손실이 발생할지
- (투자 위험의 마지노선이라 생각하면 편함)
- ✓ 수익률이 높고 MDD가 낮은 전략이 우수

Overview



퀀트 데이터와 가격 데이터 전처리 -> 딥러닝 모델 학습 후 수익률 예측 -> 상위 20개 종목 포트폴리오 구성
-> 3개월 뒤 다시 종목 재선정

데이터

1. 주식데이터

- 과거는 회귀분석 사용
- 하지만 주가의 흐름은 비선형적
- 비선형 함수로 출력되는 딥러닝 구조를 통해 예측 보완
- 현재는 추세를 나타내는 보조지표를 사용해 주가예측
- 보조지표로 ROC, RSI, MACD, 이동평균선 등

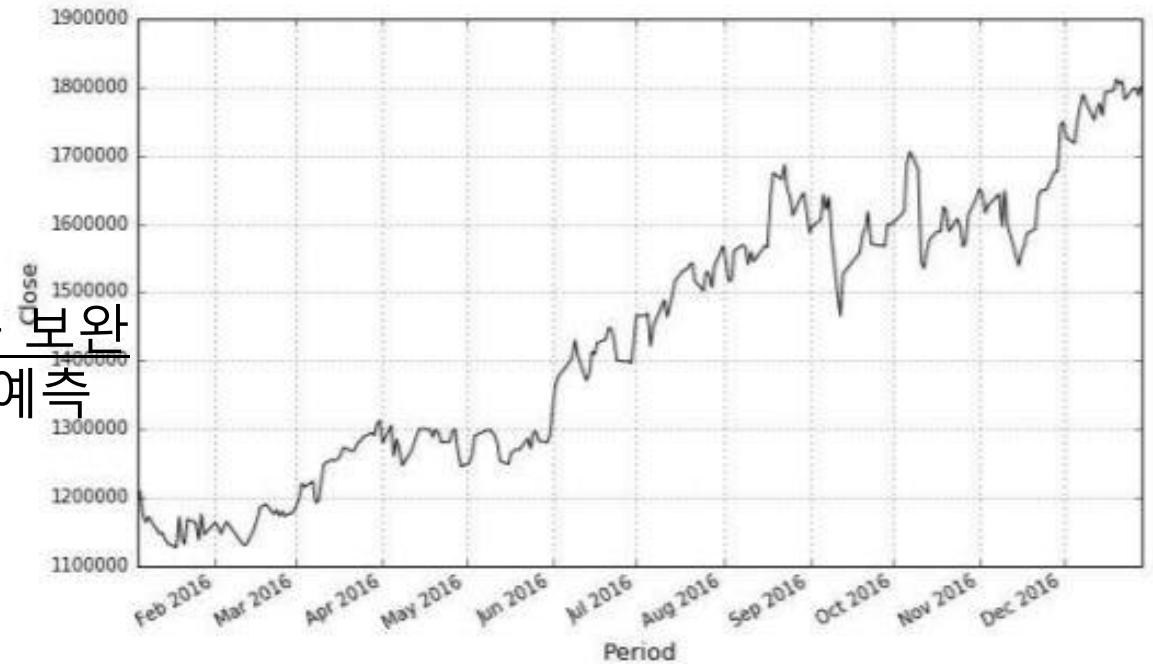
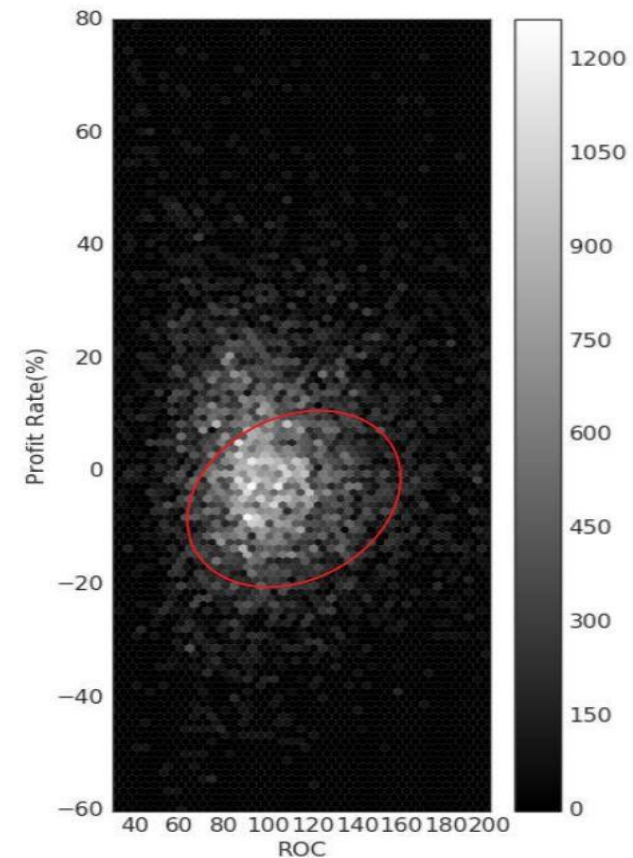


그림 2. 2016년 1월부터 12월까지의 삼성전자의 종가

1. 주식 데이터

- ROC (Price Rate Of Change) : 과거로부터 가격이 어떻게 변화했는지 나타내는 보조 지표

$$ROC = 100 \times \frac{\text{금일 종가} - n\text{일전 종가}}{n\text{일전 종가}}$$



1. 주식데이터

- MACD (Moving Average Convergence Divergence)
: MACD는 장기이동평균과 단기이동평균의 차이를 보여주는 보조 지표

분포의 형태가 사선이 아닌 수직의 형태 MACD 값에 따른 변화는 미미하다
작은 MACD 값의 변화에 따라 수익률도 같이 변함

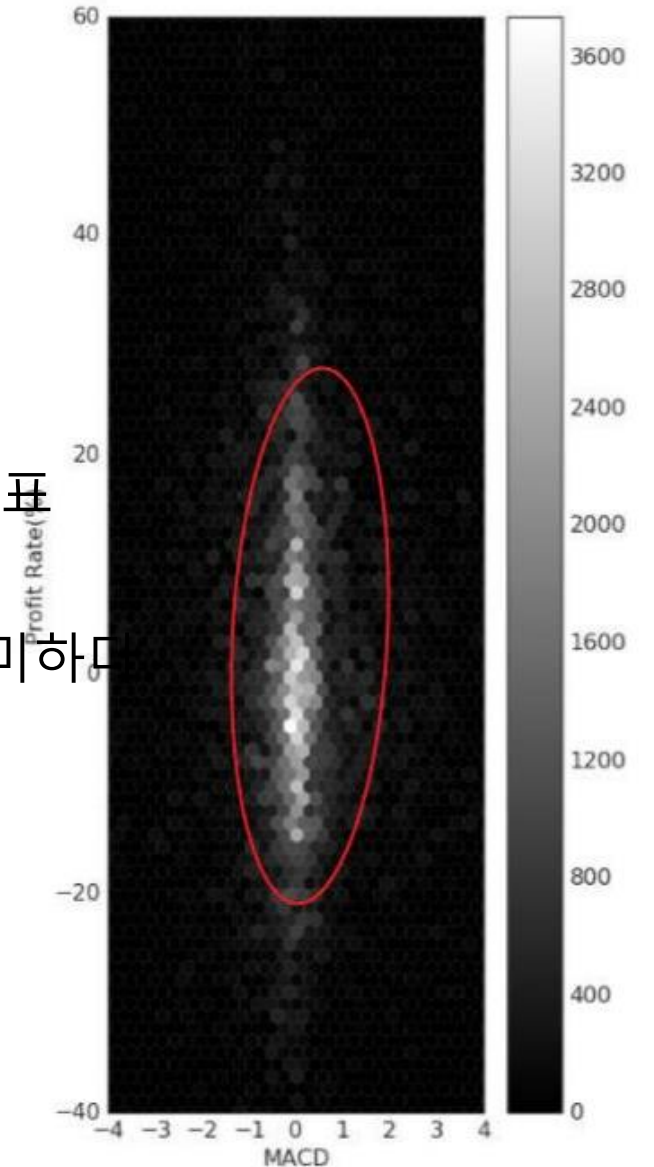


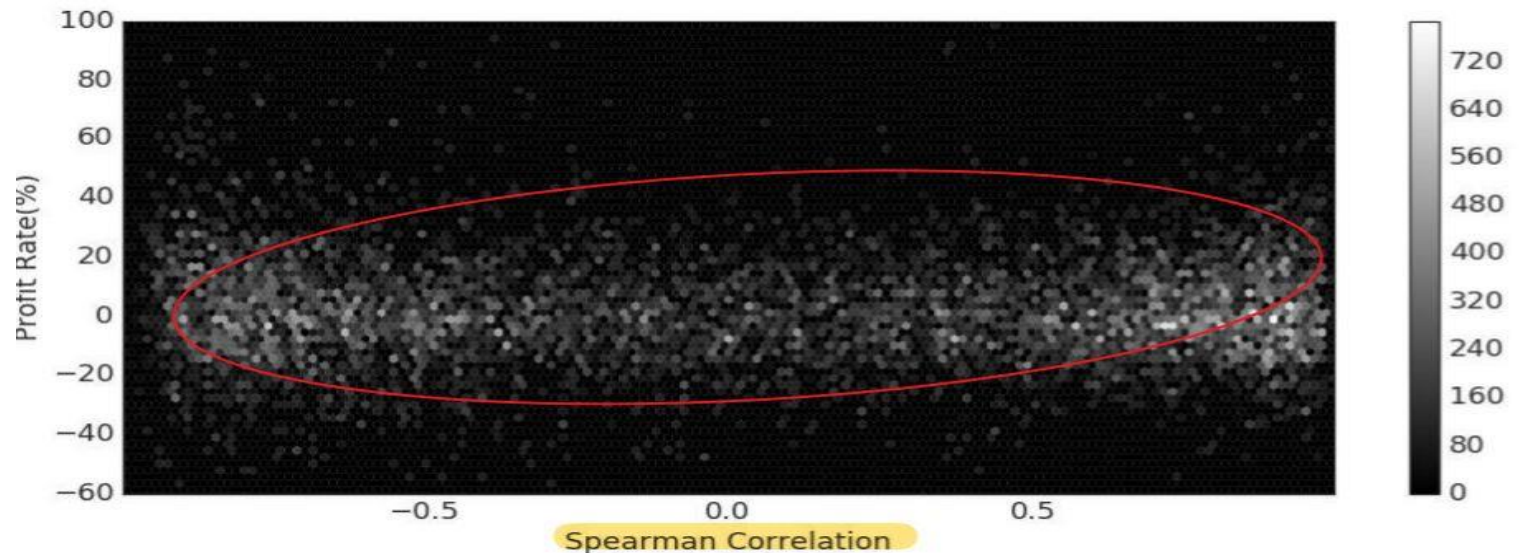
그림 4 2007년부터 2017년까지 MACD에 따른 수익률의 분포

1. 주식데이터

- 스피어만 상관계수 (Spearman Correlation)

: 두 요소에 대해 각각 랭크를 매겨 **두 요소 간의 연관성을 파악하기 위해 사용**
주식예측에 있어서는 주가가 꾸준히 상승했는지 파악하기 위해 사용

- 스피어만 상관 계수값이 **1에 가까워질수록** 혹은 **-1에 가까울수록** 높아짐
- 상관이 반대로 될 때에 수익률이 더 높은 경향이 있음



2. 퀀트 데이터

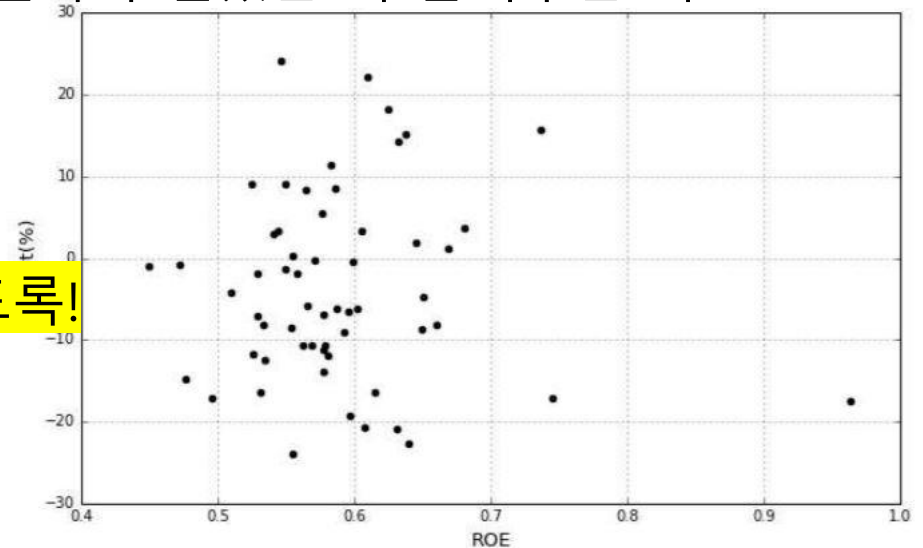
- ✓ 가치투자자는 공개된 재무제표를 통해 **기업의 미래성장가치**를 표현 할 수 있는 지표를 만들어, 지표를 토대로 거래를 한다.

지표 **가치지표** - 기업의 현재가치를 파악하는 방식 => 기업의 현재가치가 예상가치보다 낮은 경우 투자

퀄리티 지표 - 기업의 상태가 좋은 지 나타내는 방식 => 기업의 상태가 좋은, 우량주에 투자

Ex) ROE(Return On Equity)는 자본대비 순이익을 얼마나 벌었는 지 알려주는 지표

여러 지표를 동시에 입력해 미래의 수익률을 분별할 수 있도록!



자기자본 이익률

그림 6. 2016년 1분기 시가총액 상위 100종목의 ROE에 따른 3개월 뒤 수익률

수익률 예측 방법

- ✓ FNN(Feed Forward Neural Networks) 을 이용
- ✓ 주식의 상승률을 이미 알고 있으며(논문에서) 데이터로부터 수익률을 예측하는 지도학습
- ✓ 회귀모델 통해 예측 (상승 or 하락)

$$h_{i+1} = \delta(W_1 X + b_1) \quad (1)$$

$$h_{i+2} = \delta(W_2 h_{i+1} + b_2) \quad (2) \quad \text{활성화 함수}$$
$$\vdots$$

$$y = W h_{i+n} + b_L \quad (3) \quad \text{최종 출력 값}$$

$$W_{new} = W_{old} - \ell \cdot \frac{\partial E}{\partial W_{old}} \quad (5)$$

$$b_{new} = b_{old} - \ell \cdot \frac{\partial E}{\partial b_{old}} \quad (6)$$

- ⇒ 예측 값이 실제 값에 가까울 출력을 낼 수 있도록 훈련 : 역전파 (손실함수 MSE)
- Dropout, batch normalization, SeLu 사용

데이터 전처리

- 가격 데이터

- 과거로부터의 가격 추세에 대한 정보 포함
- Feature scaling 수행

- 쿼트 데이터

- 최근 분기의 데이터는 현재의 주식의 상태를 파악하는 것이 중요
- 4개 분기 feature scaling -> 현재분기의 값을 0~1 사이의 값으로 줄임
- 특히 지표의 변화에 대해 중점을 가지고 분석

Deep Neural Networks 를 사용한 중장기 투자 전략

- ✓ DNN을 이용한 중장기 투자는 포트폴리오 전략을 이용
- ✓ 포트폴리오 전략은 수익률과 MDD 이 평가 요소
- ✓ 즉 **좋은 중장기 투자전략은 수익률이 높고 MDD가 낮아야 함**

$$\text{수익률} = 100 \times \left(\frac{\text{현재 자본}}{\text{초기 투자금}} - 1 \right)$$

$$\text{MDD} = 100 \times \left(\frac{\text{최대자본 갱신 이후 발생한 최소 자본}}{\text{최대 자본}} - 1 \right)$$

- ✓ 포트폴리오 전략은 보유할 종목 숫자를 정해야 함
- ✓ 종목 수가 많을 시 고수익 확률이 낮아지고, 종목 수가 적을 시 고수익 확률이 높아짐
- ✓ 논문에서는 **상위 20종목을 포트폴리오로 구성**
- ✓ 이 20종목이 실제 상승률이 높은 종목과 일치했는 지 파악이 중요!
- ✓ 이를 위해 포트폴리오의 **구성유사도와 스피어만 상관계수** 사용

실험 결과

총 3가지 모델 평가

1. 주식의 가격데이터만 훈련한 모델
2. 퀀트데이터만 훈련한 모델
3. 퀀트와 주가데이터를 함께 훈련한 모델

- ✓ 각각 데이터 훈련한 모델이 더 우수하다
- ✓ 포트폴리오 유사도가 높다 해서 수익률이 높진 않
-> 전체적인 경향에 대한 예측이 수익률 보장X

Year	Quarter	Model	PS(%)	KOSPI100	top20	top10
2015	1Q	Price	20	-0.009	-0.089	-0.039
		Quant	30	0.131	-0.086	0.152
		Price + Quant	20	-0.114	0.060	0.273
	2Q	Price	35	0.083	0.221	0.118
		Quant	35	-0.042	-0.099	0.188
		Price + Quant	20	-0.032	0.200	0.564
	3Q	Price	35	-0.056	0.128	-0.015
		Quant	15	-0.131	-0.361	-0.406
		Price + Quant	10	-0.089	0.2676	0.478
	4Q	Price	40	0.173	0.162	0.031
		Quant	35	0.077	-0.077	0.406
		Price + Quant	35	0.040	-0.053	0.006

2016	1Q	Price	10	0.002	0.027	0.129
		Quant	15	-0.088	-0.068	-0.455
		Price + Quant	20	0.192	-0.1413	0.0787
	2Q	Price	30	0.010	0.107	0.062
		Quant	25	0.069	0.245	0.430
		Price + Quant	20	-0.030	0.032	-0.261
	3Q	Price	20	-0.079	0.169	-0.063
		Quant	30	-0.037	-0.015	0.042
		Price + Quant	35	0.003	-0.009	0.115
	4Q	Price	15	0.022	0.643	-0.068
		Quant	25	-0.086	-0.088	0.394

모의결과 결과- 기간에 따른 수익률과 MDD 파악이 목표

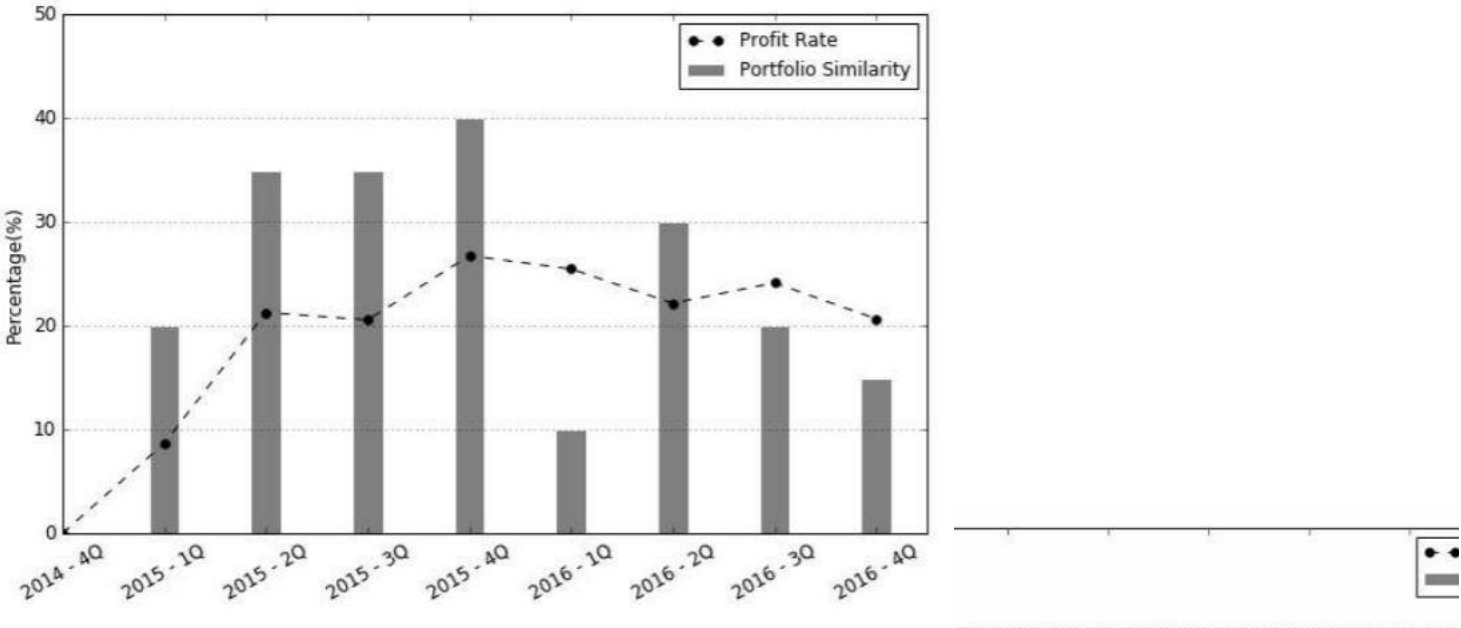


그림 12 가격모델의 분기별 수익률과 포트폴리오 유사도

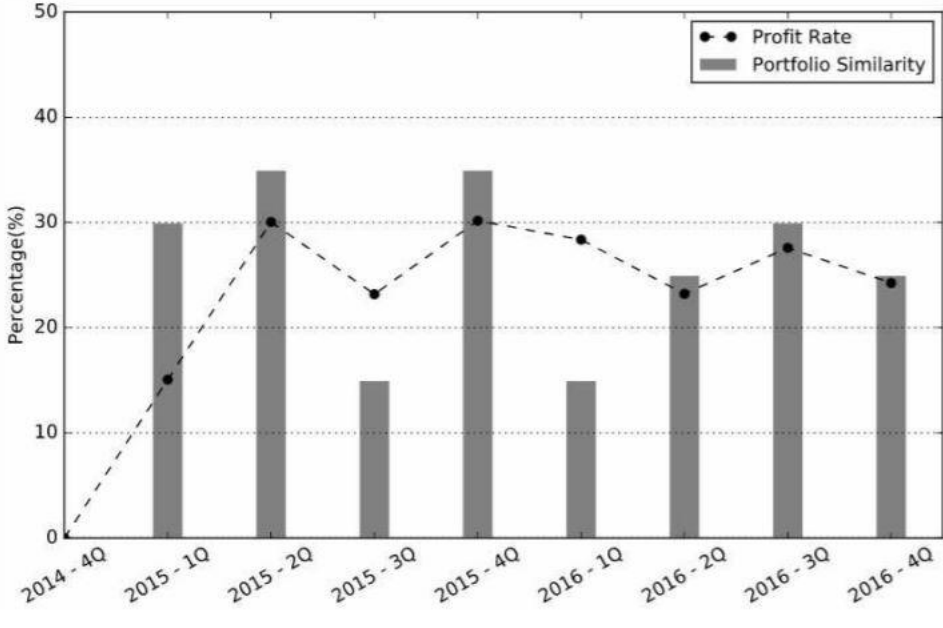


그림 13 퀀트모델의 분기별 수익률과 포트폴리오 유사도

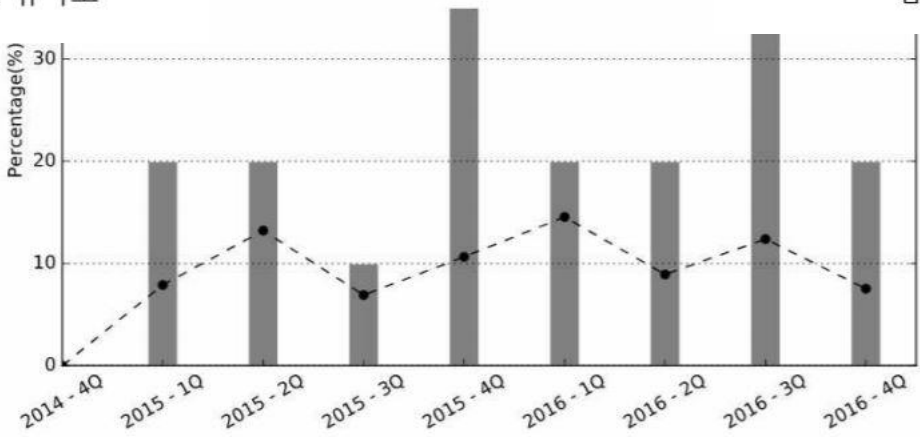


그림 14 가격+퀀트 모델의 분기별 수익률과 포트폴리오 유사도

모의결과 결과 - 기간에 따른 수익률과 MDD 파악이 목표

✓ 각각 데이터 학습한 결과가 좋음

-> 가격데이터와 퀀트 데이터가 예측하는 특성이 다르기 때문

✓ 가격모델의 경우 퀀트 데이터 보다 수익률이 낮지만 수익률의 변동폭이 크지 않음 -> 안전한 수익

✓ 포트폴리오 유사도와 수익률이 반드시 일치하지 않지만 유사도가 높을 수록 하락에도 잘 견딤

✓ 최적 모델은 퀀트데이터로만 학습한 모델!

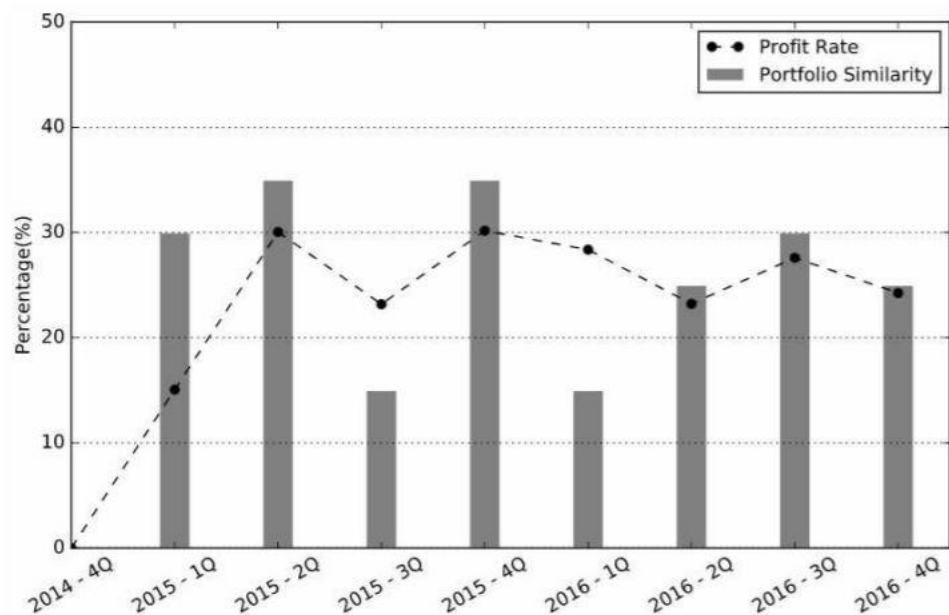


그림 13 퀀트모델의 분기별 수익률과 포트폴리오 유사도

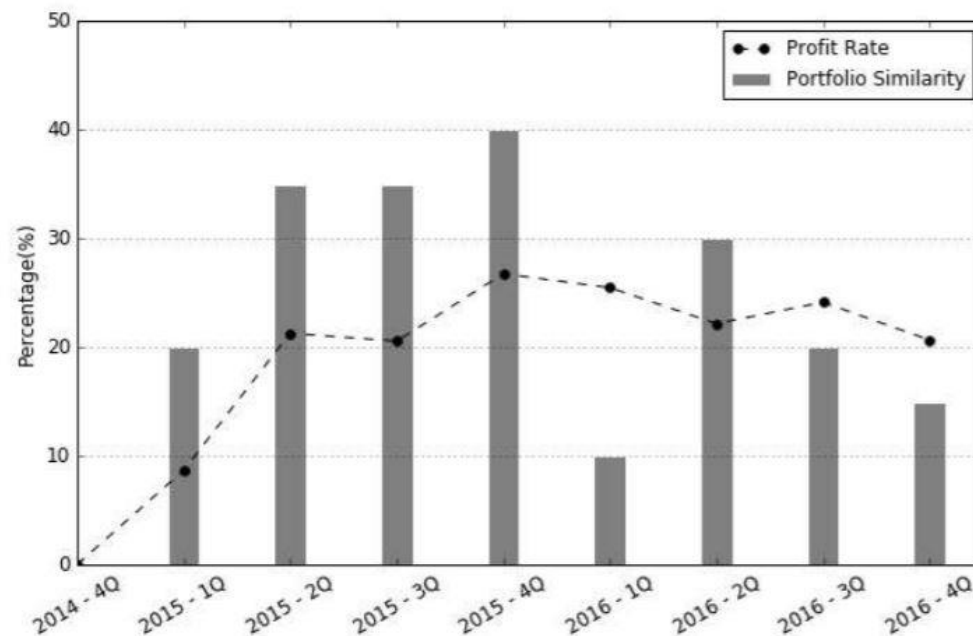


그림 12 가격모델의 분기별 수익률과 포트폴리오 유사도