



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

석사학위 논문

지도교수 황 찬 규

퀀트분석을 이용한 포트폴리오
구성에서 Monte Carlo Simulation을
이용한 DQN 강화학습 연구

A Study on DQN Reinforcement Learning
Using Monte Carlo Simulation in Portfolio
Configuration Using Quant Analysis

2020

서울벤처대학원대학교

융합산업학과 U-City · 융합기술경영전공

김 의 철

퀀트분석을 이용한 포트폴리오
구성에서 Monte Carlo Simulation을
이용한 DQN 강화학습 연구

A Study on DQN Reinforcement Learning
Using Monte Carlo Simulation in Portfolio
Configuration Using Quant Analysis

이 논문을 석사학위 논문으로 제출함.

2021년 1월

서울벤처대학원대학교

융합산업학과 U-City · 융합기술경영전공

김 의 철

김의철의 석사학위

논문을 합격으로 판정함.

심사위원장 김지형 인

심사위원 도명정 인

심사위원 함관기 인

2020년 12월

서울벤처대학원대학교

연구윤리서약서

본인은 서울벤처대학원대학교 대학원생으로서 이 학위논문 작성과정에서 다음과 같은 연구윤리 기본원칙의 준수를 서약하고 이를 지켰음을 확인합니다.

첫째, 지도교수의 지도에 따라 정직하고 엄정하게 연구를 수행하여 학위논문을 작성한다.

둘째, 논문작성 시 타인에의 대필의뢰, 타 논문의 표절, 데이터의 위조와 변조 등 학문적 진실성을 훼손하는 어떠한 연구 부정행위도 하지 않는다.

2020년 12월

학 위 명 : 경영학석사

학 과 : 융합산업학과

U-City · 융합기술경영

성 명 : 김 의 철



서울벤처대학원대학교 총장 귀하

감사의 글

언제 어디서 무엇이 왜 일어나는지 사람은 누구나 알고 싶어 합니다.

그래서 어떤 사람은 비과학적 방법으로 어떤 사람은 과학적 방법으로 해답을 찾고자 합니다.

그 안에는 욕망과 두려움이 있습니다.

그러나 해답은 찾는 것이 아니고 주어지는 것이라는 것을 찾아가는 과정에서 항상 깨달아 갑니다.

하느님의 뜻에 따라 이끄심을 믿으며, 감사와 찬양을 드립니다.

항상 새로운 도전과 정진에 길을 열어주시고 지도해 주시는 황찬규 교수님께 감사를 드립니다.

따뜻한 시선과 격려를 전해 주신 선-후배 동기 여러분께도 감사드립니다.

형제들과 조카들에게 한결같은 사랑을 전합니다.

성실과 인내의 사랑을 가르쳐 주시고 먼저 하늘나라에 가신 아버지 그리고 배우고 나누는 사랑을 가르쳐 주시는 어머니 사랑합니다.

2021.1

김 의 철 올림

목 차

표 목차	iii
그림 목차	iv
수식 목차	vii
논문 개요	ix
I. 서 론	1
1. 연구의 필요성 및 목적	1
2. 연구방법 및 논문의 구성	6
II. 이론적 배경	12
1. 고전적 금융 시계열분석	12
2. GARCH 모형	14
3. RNN(LSTM) 모형	18
4. DQN강화학습	22
5. GBM Monte Carlo Simulation	26
6. RNN, 강화학습 포트폴리오 구성 선행연구	28
III. 종목선정 및 방법 적용	33
1. 데이터 웹 크롤링	33
2. 퀀트분석을 이용한 종목선정	36

3. 변동성 분석과 포트폴리오 구성	56
4. GBM을 이용한 DQN	75
IV. 백 테스트 및 분석	84
1. 백 테스트 환경설정	84
2. 테스트 결과 및 분석	85
V. 결 론	96
1. 연구결과 요약 및 시사점	96
2. 연구한계 및 향후 연구방향	98
참고문헌	102
ABSTRACT	108

표 목 차

<표 3-1> 최소변동성 기준 10종목의 비중	58
<표 3-2> 최소변동성 기준 10종목의 비중(비중 제약조건)	59
<표 3-3> 최대변동성 기준 10종목의 비중	60
<표 3-4> 위험균형 기여위험률 평준화 10종목의 비중	61
<표 3-5> X042670(두산인프라코어) 종목 GARCH 모형의 계수	66
<표 3-6> X042670(두산인프라코어) 종목 GARCH 모형의 유의성	66
<표 3-7> X042670(두산인프라코어) 종목 GARCH 모형의 Ljung-Box test	67



그 립 목 차

[그림 1-1] 연구의 구성도	9
[그림 1-2] 분석과정 순서도	10
[그림 1-2] 분석과정 순서도	11
[그림 2-1] RNN의 알고리즘 작동	19
[그림 2-2] RNN과 LSTM 도식과 수식	21
[그림 2-3] 강화학습 개념도	22
[그림 2-4] Bellman Equation	24
[그림 2-5] DQN 알고리즘	26
[그림 2-6] GBM의 간략 도식화	27
[그림 3-1] 전체시장 WICS 분류	34
[그림 3-2] 순위 종목 누적 시가총액 합계	35
[그림 3-3] KOSPI 200 ETF 2년간 가격변동과 연간 수익률	37
[그림 3-3] KOSPI 200 ETF 2년간 가격변동과 연간 수익률	38
[그림 3-4] KOSPI 200 ETF와 키움증권 및 포스코강판 베타(β)선	39
[그림 3-5] 주식시장 산업군 임의 선정 기업 Beta(β)선	40
[그림 3-6] 주가종목 수익률 변동성 히스토그램	41
[그림 3-7] 수익률 변동성 하위 30위 히스토그램	42
[그림 3-8] 모멘텀(위험조정) 수익률 30위 주가 그래프	44
[그림 3-9] 모멘텀(위험조정) 수익률 30위 산업분류별	44
[그림 3-10] 종목선정 투자가치 지표 간의 상관계수	46
[그림 3-11] 종목선정 재무지표 간의 상관계수	47
[그림 3-12] PER, PBR, ROE의 주식시장에서의 분포	49

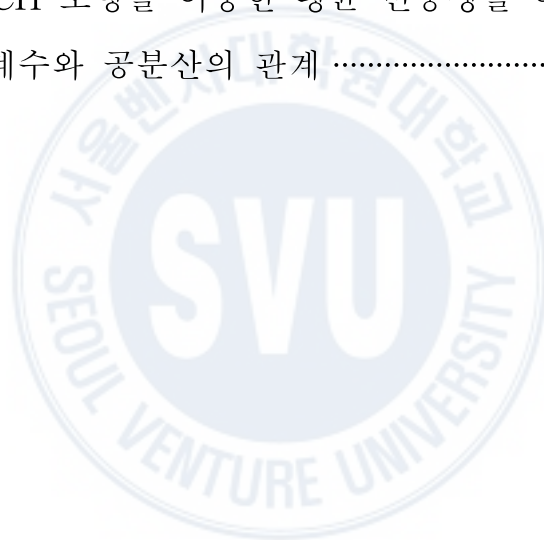
[그림 3-13] F-SOCRE에 따른 주식시장 분포	51
[그림 3-14] 종목선정 지표의 시장종목 순위 분포	52
[그림 3-14] 종목선정 지표의 시장종목 순위 분포	53
[그림 3-15] 종목선정 지표 간의 상관계수	54
[그림 3-16] 선정 10종목의 주가변동 그래프	55
[그림 3-17] 선정 10개 종목의 상관계수	57
[그림 3-18] 최소변동성 기준 10종목의 비중	59
[그림 3-19] 최소변동성 기준 10종목의 비중(비중 제약조건)	60
[그림 3-20] 최대변동성 기준 10종목의 비중	61
[그림 3-21] 위험균형 기여위험률 평준화 10종목의 비중	62
[그림 3-22] 10종목에 대한 log 수익률과 ACF 그래프	64
[그림 3-23] X042670(두산인프라코어) 종목에 대한 log 수익률과 GARCH 모델	65
[그림 3-24] 10개 종목에 대한 log 수익률과 GARCH 모델	69
[그림 3-25] X042670(두산인프라코어) 종목에 대한 log 수익률과 종가	71
[그림 3-26] X042670(두산인프라코어) 종목에 대한 1 Step의 예측 2개	72
[그림 3-27] X042670(두산인프라코어) 종목에 대한 LSTM 모델 학습의 Loss 변화 ·	72
[그림 3-28] X042670(두산인프라코어) 종목에 대한 LSTM 모델의 예측	73
[그림 3-29] X042670(두산인프라코어) 종목에 대한 다변량 LSTM 모델의 예측	74
[그림 3-29] X042670(두산인프라코어) 종목에 대한 다변량 LSTM 모델의 예측	75
[그림 3-30] X042670(두산인프라코어) 종목의 1 step의 GBM MC simulation 100회 생성	76
[그림 3-31] 전체 데이터 훈련의 훈련과정 확인 투자 100회	78
[그림 3-32] 63일 Simulation 데이터 훈련과정 확인 투자 100회	79
[그림 3-33] 252일 Simulation 데이터 훈련과정 확인 투자 100회	80
[그림 3-34] 252일 Simulation 데이터 1500회 훈련 확인 투자 100회	81

[그림 3-35] 백 테스트에 사용된 3종목의 주가	84
[그림 3-35] 백 테스트에 사용된 3종목의 주가	85
[그림 3-36] 최소변동성 기준 동적 비중 전략	86
[그림 3-37] GARCH를 이용한 변동성 예측	87
[그림 3-38] GARCH를 이용한 최소변동성 기준 동적 비중 전략	88
[그림 3-39] RNN을 이용한 수익 변동 예측	89
[그림 3-39] RNN을 이용한 수익 변동 예측	90
[그림 3-40] RNN을 이용한 최소변동성 기준 동적 비중 전략	91
[그림 3-41] 전체 데이터 DQN 강화학습 포트폴리오 테스트 투자 100회	92
[그림 3-42] GBM-MC(63일) DQN 강화학습 포트폴리오 테스트 투자 100회	93
[그림 3-43] GBM-MC(252일) DQN 강화학습 포트폴리오 테스트 투자 100회	94

수식 목차

(수식 1) 자본시장선(CML)	12
(수식 2) 증권시장선(SML)	12
(수식 3) 체계적 위험: 베타(β)	13
(수식 4) 지수평활(simple exponential smoothing)	15
(수식 5) 자기회귀 모델(auto regressive models)	16
(수식 6) 이동평균 모델(Moving Average)	17
(수식 7) GARCH 모델	18
(수식 8) GARCH(1, 1) 모델	18
(수식 9) 상태 전이 확률(State transition probability matrix)	23
(수식 10) 보상(reward)	23
(수식 11) 정책(policy)	23
(수식 12) Q-value 함수	25
(수식 13) GBM Monte Carlo Simulation	27
(수식 14) GBM MCS 랜덤한 방향성	28
(수식 15) 샤프지수(sharp 비율)	43
(수식 16) PBR(Price to Book-value Ratio:주가순자산비율)	45
(수식 17) PER(Price to Earnings Ratio:주가순이익비율)	45
(수식 18) PCR(Price to Cashflow Ratio:주가현금흐름비율)	45
(수식 19) PSR(Price to Sales Ratio: 주가매출비율)	45
(수식 20) ROE(Return On Equity:자기자본이익률)	47
(수식 21) GP/A	47
(수식 22) CFA	47

(수식 23) 공분산(Covariance)	56
(수식 24) 공분산 행렬식 형태	56
(수식 25) 위험기여도(MRC: Marginal Risk Contribution)	56
(수식 26) 위험기여도 행렬식 형태	56
(수식 27) ARIMA 검정 통계량 Q	63
(수식 28) LM(Ljung-Box test)	63
(수식 29) GARCH(1, 1) 모형	65
(수식 30) GARCH 모형을 이용한 평균 변동성을 예측	68
(수식 31) 상관계수와 공분산의 관계	70



논 문 개 요

4차 산업혁명 시대 그 바탕에는 컴퓨터 성능 및 네트워크의 발달과 함께 지능형(스마트) 알고리즘의 발전이 배후에 있다. 이러한 현상을 대표하는 단어 SMART란 기존의 컴퓨터 계산 능력을 유-무선으로 연결함을 넘어, 인간의 판단력과 분별력이 필요한 업무에 자동화 알고리즘을 통하여 상황에 따라 즉각적으로 인간의 역할을 보완 또는 대체하는 지능(Intelligence)형 시스템을 탑재한 상태이다. SMART가 가능한 바탕에는 통계분석과 수학적 기법의 발전이 있다. 그리고 이러한 방법론들은 빅데이터, IOT, AI 등의 컴퓨터 통신환경과 더불어 연속적인 시계열 데이터들을 분석하고 활용하는 방향으로 발전하고 있다.

본 연구에서는 이러한 발전 방향에 있어서 금융분석과 투자 분야에서 사용하는 대표적 시계열분석 방법을 개괄하고 심층학습(Deep Learning), 강화학습과 같은 기계학습(Machine Learning) 기법들과 함께 실제 주식투자 포트폴리오 분석에 적용해 보았다. 더 나아가 실제 데이터가 아닌 가상의 데이터를 생성하여 훈련한 강화학습 모델이 어느 정도 주식투자 포트폴리오에 대하여 학습 가능한지 살펴보았다.

첫째, 이를 위하여 먼저, 주가 포트폴리오 투자와 시계열분석에 대한 이론적 배경으로 자본자산가격결정모형(CAPM), 일반화자기상관조건부이분산성(GARCH)모형, 순환신경망(RNN)을 살피고 마지막으로 DQN 강화학습과 GBM Monte Carlo Simulation에 대하여 살펴보았다.

이러한 방법론을 모델로 구현하기 위하여, 웹 크롤링 방법을 사용하여 주식 종목에 대한 재무제표, 주가 등의 정보들 수집하였다. 이후 이를 정리하여 가치지표, 재무지표, 모멘텀 지표로 정리하고 포트폴리오의

구성 종목선정에 사용하였다.

둘째, 선정된 주가 종목을 시계열분석 방법을 이용하여 포트폴리오 구성비를 계산하는 데 사용해 보았다. 이때 포트폴리오 구성 전략은 최소변동성 비율 전략을 같은 전략 조건으로 설정하고 4가지 방법론들의 분석적 특징을 실제 주가 데이터에서 살펴보고자 하였다.

이중 DQN 강화학습은 실제 데이터뿐만 아니라 GBM Monte Carlo Simulation을 사용하여 가상의 데이터를 2가지 만들어 별도로 학습을 시켜 가상의 데이터를 통한 학습 가능성을 살펴보았다.

셋째, 비교 테스트를 위하여 4가지 방법론으로 준비된 모델들을 다른 패턴을 보이는 3개 종목의 주가 데이터에 백 테스트(Back Test)를 진행하였다. 각기 다른 방법으로 구성된 모델로부터 포트폴리오를 구성하고 그 누적수익률 결과를 비교 분석하였다.

이러한 비교를 통하여 분석기법에 사용된 지능형 알고리즘의 발전 방향이 정형화되어 기록된 데이터 수치를 계산하는 모델에서 지속해서 발전하고 있음을 확인하고자 하였다. 먼저, 환경 모형을 설정하고 입력 변수로부터 결과를 예측하는 방법으로 발전하였고, 다시 데이터 자체로부터 관계성과 특징을 찾아내는 기계학습(Machine Learning) 방법으로 발전하고 있는 것을 알 수 있었다.

넷째, 이러한 SMART 방법론의 발전이 주어진 데이터의 한계를 넘어서 환경을 반영하는 알고리즘을 도입한다면 예측 또는 가능성 있는 변화 상황을 SMART 알고리즘 스스로가 학습해 나갈 수 있을지 그 가능성에 대하여 살펴보고자 하였다.

이러한 방법론을 적용하고 비교·분석한 본연구의 결과는 다음과 같다. 먼저, 시계열분석의 모델들의 방법을 개괄하고 금융분석에 적용해 본

결과 분석방법은 환경적 변화 요인에 대하여 적극적으로 예측 반영하는 방향으로 발전하고 있음을 알 수 있었다. 다음으로 각 방법으로부터 생성한 모델들을 실제 주가 데이터에 적용해 포트폴리오를 구성해 누적수익률을 비교해 본 결과 분석방법이 발전할수록 데이터의 변화에 대하여 모델이 예측 반응하는 결과를 보여주었다. 마지막으로 강화학습을 이용하여 주가 데이터에 대한 상황판단과 그에 따른 포트폴리오 구성비에 대한 정책을 학습시킨 모델을 실제 주가 데이터에 100회 적용한 결과, 손실이 발생하는 경우가 있었으나, 대다수의 경우 누적수익률이 수익을 발생시킴을 알 수 있었다.

여기서 더 나아가 가상의 데이터를 이용하여 학습시킨 모델에서도 실제 데이터만을 사용하여 학습시킨 모델과 같이 수익이 발생하는 경우가 증가하였다. 다만 가상의 데이터의 길이가 길어질수록 더 높은 수익률을 발생시키는 경우가 발생하는 반면 더 손실을 발생시키는 경우도 발생하여 안정성이 떨어짐을 알 수 있었다. 이러한 문제는 추가적인 훈련을 통하여 안정할 수 있을 것으로 보인다. 실제 주가 데이터를 수집하는 비용과 노력에 비교하여 더 빠르고 효율적 학습이 가능할 것으로 보인다.

중요한 결과는 주가의 무작위 변화를 알고리즘으로 발생시키고 이를 통해 SMART 알고리즘이 학습될 수 있음을 확인한 점이다. 이는 투자 환경에 변화 요인들을 역사적 데이터로 분석하는 것이 아니라 가능성에 대한 알고리즘으로 입력하고 이에 대비한 정책을 SMART 시스템으로 학습시킬 수 있음을 뜻한다.

I. 서론

1. 연구의 필요성 및 목적

가. 시계열 데이터 분석

현재의 사회는 4차 산업혁명의 시대이다. 그 바탕에는 컴퓨터 성능의 발달과 네트워크를 넘어 지능형(스마트) 알고리즘의 발전이 배후에 있다. 1990년대는 컴퓨터의 유선 네트워크의 발달 시대였고 2000년대는 무선 모바일 네트워크의 시대, 2010년대는 빅데이터의 시대였다. 2020년대부터 놀랍도록 발전하고 있는 SMART는 앞으로는 급속하게 발전하여 다양한 분야에서 서비스와 부가가치를 만들어낼 수 있다고 예측한다. 이처럼 SMART의 정의는 분야별, 사용자별로 다양하게 정의되고 있다. 본 연구에서는 SMART의 정의를 지능(Intelligence)이라는 단어에 집중해 해석하고자 한다. SMART란, 기존의 컴퓨터 계산 능력을 유-무선으로 연결하는 것에서 더 나아가, 인간의 판단력과 분별력이 필요한 업무를 자동화 알고리즘을 통하여 상황에 따라 즉각적으로 보완 또는 대체하는 시스템을 탑재한 상태로 정의한다. 이러한 SMART 시대는 IOT 등의 상품시장에서부터 생산과정의 관리, 시장의 분석까지 모두 분야에서 영향을 받을 것이다. 따라서 이러한 SMART의 핵심인 지능(intelligence)화를 가능하게 하는 알고리즘을 알아가는 것은 중요하다. 특히, 모든 정보와 데이터가 네트워크를 통해 연결되어 시간의 변화에 따라 지속 제공되고 배포되는 현대 사회에서 연속적인 현상을 분석하는 알고리즘의 이해는 앞으로 더욱 중요해질 것이다.

이러한 알고리즘의 바탕에는 통계분석과 수학 기법이 있다. 기존의

통계기법이 평균(기댓값)의 차이 분석 및 집단을 분류하는 방법에서 분산분석과 같은 집단 간 차이를 분석하는 수준으로 발전하였고 더 나아가 교차분석과 요인분석이 더해지면서 인과관계를 구하는 회귀분석으로 발전해왔다. 거기서 다시 복잡한 인과관계의 구조를 구조모형을 통하여 또는 다양한 회귀분석 방법이나 기계학습(Machine Learning)기법을 통하여 설명하고자 하였다. 그리고 분석 모델이 시간의 흐름 또는 작동의 흐름에 따라 어떻게 변화하고 작동할 것인지 이해하고 예측하려는 방향으로 진행되고 있다.

이러한 발전에도 현재 금융분석 분야는 절대적 분석방법이 나오지 않고 있다. 그 이유는 무한하게 변화 가능한 환경변수에 대한 입력의 한계와 예측할 수 없는 급격한 변동에서 오는 문제라고 할 수 있다. 즉, 자연현상의 복잡함에 인간의 복잡한 심리와 행동 관계가 더해지면서 특정 몇 가지 현상을 해석하는 수 개의 모델로는 충분한 해석이 되지 않기 때문이다. 어떠한 경제 상황이 발생할지는 누구도 정확히 예측할 수 없다. 이렇듯 어려운 분야지만 또한 인간의 물질에 대한 욕망만큼 끝없이 연구되는 분야이기도 하다.

이렇듯 거의 불가능해 보이는 예측이지만 투자 측면에서 보면 현상을 해석하는 것이 아니라 최선의 대응 행동정책을 세우기 위한 수단으로 분석법을 통해서 얻을 수 있는 것이 많다. 그것은 투자행위의 목적이 현상의 이해와 해석이 아니라 행위의 목적 즉, 수익에 있기 때문이다. 어느 정도 신뢰 범위 내의 현재 상황의 이해와 예측을 통해서 가장 좋은 결과를 가져올 행위 정책을 준비할 수 있기 때문이다.

이러한 측면에서 과거 경제학적인 연구방법들이 현상을 해석하는데 주목하고 있다면 최근 연구되고 있는 SMART 알고리즘 연구들은 예측

과 예측을 이용한 행동정책 수립에 집중되고 있다. 변화하는 수치를 예측하고자 하는 관점의 최근 선행연구에는 박정빈(2019)의 RNN을 이용하여 수익률의 고저를 예측하는 연구, 김기범의(2018) KOSPI200 선물 가격변화 예측, 황승환(2020)의 원/달러 환율예측력 분석을 통한 기존 시계열 예측과 LSTM모델의 비교 등이 있다. 예측과 투자정책을 같이 연구한 연구로는 김주봉(2019)의 A3C 강화학습으로 포트폴리오의 구성 정책을 학습하는 연구가 있으며, 투자정책 방향을 강화학습과 이산행동 분석 전문가모형을 앙상블학습으로 연구한 임준범(2019)의 연구가 있다.

현상의 분석과 해석을 주가 데이터의 시계열분석 기법에 관한 연구로 본다면 행위 정책은 포트폴리오의 구성에 관한 연구라고 볼 수 있다. 이런 관점에서 볼 때 SMART란 개념 정의가 금융투자에 어떻게 도입되는지 알 수 있다. 주식시장 상황의 분석과 그에 대한 포트폴리오 구성 행위에 필요한 인간의 분별력과 판단력을 보완하거나 대체하는 알고리즘이 도입되는 것이다.

본 연구에서는 이러한 알고리즘을 개괄해서 살펴보고 방법론을 주식 분석과 포트폴리오 구성에 적용해 보았다. 이후 상황의 이해와 행위 정책을 하나의 과정으로 학습하는 강화학습을 구현해 보고자 하였다. SMART 방법을 포트폴리오 구성에 직접 구현하여 인공지능이 인간의 분별능력과 판별 능력을 대체하는 자동화 수준을 얼마나 발전시킬지 그 가능성을 이해하고자 하였다.

본 연구는 이러한 이해를 위해서 황승환(2020)의 연구에서와같이 각 방법론의 비교와 함께 김주봉(2019)와 임준범(2019)의 연구와 같이 투자 정책의 자동화의 가능성을 보기 위하여 2020년 7월 21일 기준일 이전 2년간의 실제 주가 데이터로 모델을 만들고 기준일 이후 2개월간의 백

테스트를 위한 실제 주가 데이터를 사용하여 실증 비교하였다.

나. Simulation 데이터와 강화학습

기계학습(Machine Learning)은 알고리즘이 충분히 학습 훈련할 수 있는 데이터가 필요하다. 그러나 모든 상황에서 충분한 데이터가 있는 것은 아니다. 이는 주식시장에서도 마찬가지이다. 기업마다 주식시장에 진입한 기간의 길이가 다르고 늦게 진입한 경우 그 기간이 짧을 수밖에 없다. 더군다나 코스닥(KOSDAQ) 시장의 기업들은 대부분 그 주가 데이터의 기간이 수년에 불과하다. 이로 인하여 훈련데이터 기간 안에 존재하지 않았던 새로운 범위의 데이터나 데이터의 패턴이 발생하였을 때, 모델의 분석력이나 예측력은 떨어지게 된다.

이러한 실측 데이터의 부족에 대하여 이미 과학과 공학 분야에서는 발생 가능한 환경변수를 무작위로 발생시켜 연구대상의 알고리즘이나 역학관계를 검증하는 모의실험 방법이 사용되었다. 그 한 가지 방법이 Monte Carlo Simulation이다. 이러한 방법을 이미 금융공학에서도 사용하여왔다. 그러나 이는 특정 투자환경에 대한 분석 모델의 평가를 위한 방법으로 사용됐다. 과거의 데이터로 검증되어 정형화된 모델에서는 새로운 데이터의 입력은 기존 분석 모델을 수정하는 의미가 있을 뿐이다. 그에 따른 새로운 투자전략 방안을 마련하는 것은 별개의 과정이다. 이는 기존 방법론이 정보의 입력으로부터 상황을 평가하는 과정과 그 평가로부터 취할 수 있는 행동에 대한 정책을 세우는 과정이 다른 과정이기 때문이다.

하지만 실제 인간의 분별과 판단 능력은 이러한 과제를 거의 동시적으로 시행한다. 논리적인 과정 해석이나 주어진 상황에 따라서는 분석

과 대응전략 검토라는 순서로 나타나기도 하지만 실제로 인간의 투자과정은 새로운 정보의 입력에 대하여 분석과 그에 대한 전략방안을 병렬에 가까운 분석, 대입 반복 과정으로 동시에 고려한다. 이런 능력을 통하여 최선의 대안을 찾아간다. 그리고 이러한 과정을 인간의 지적능력은 학습을 통하여 더 높은 결과와 효율로 수행할 수 있다.

이와 같은 인간의 능력을 모방한 기계학습 알고리즘이 강화학습(Reinforcement Learning)이다. 강화학습은 기계학습(machine learning) 방법론 중 환경의 평가와 정책의 탐색을 동시에 하는 최적화 과정을 가지고 있다. 이는 드론이나 로봇이 주어진 환경에서 목적을 수행하기 위하여 일련의 행동과정을 자동으로 선택하는 알고리즘에 사용된다.

새로운 입력 정보로부터 목표를 수행하기 위한 일련의 과정 중, 최선의 작동을 자동으로 선택하게 하고, 그 결과를 반복적으로 학습하여 목적을 최대화 또는 효율을 극대화하는 것이다.

여기에 강화학습은 실제 데이터가 아닌 가상의 Simulation 정보를 사용하여 상황판단과 행동의 정책을 동시에 학습해 갈 수 있음이 입증되고 있다. 대표적인 예로 딥-마인드사의 알파고(Alpha-GO)가 바둑게임에서 강화학습 방법을 통하여 실제의 대국 데이터뿐만 아니라 가상의 자체 대국 Simulation을 통하여 학습될 수 있음을 잘 보여주었다.

SMART 알고리즘 기법의 이러한 가능성에 대하여 강문주(2018)는 RNN을 이용한 예측 데이터와 가우시안 프로세스 모델을 이용한 가상 데이터를 실제 데이터와 결합하여 강화학습 데이터로 사용할 경우의 성능 향상을 연구하였다. 본 연구는 여기서 한 걸음 더 나아가 일정 기간의 데이터 전체를 가상의 데이터로 SMART 알고리즘 학습에 사용 가능한지 연구하였다. 그 가능성을 알아보기 위하여 충분한 기간의 데이

터를 가진 종목을 본래 실제 데이터를 가지고 강화학습 모델 학습시키는 경우와 일정 부분을 Simulation 정보로 재생성하고 이들을 강화학습 모델에 학습시키고 백 테스트 결과를 비교하였다.

본 연구는 Simulation 데이터 학습이 가능하다면 경영환경이나 투자 환경에서 SMART 알고리즘이 적용될 수 있는 범위가 환경변수의 조작적 정형화에 따라 확대될 수 있다고 보고 있다. 환경변수를 조작적으로 입력 정보화할 수 있는 만큼 실제 데이터가 아닌 가상의 데이터를 통하여 강화학습은 그 변수에 대하여 적응한 전략을 학습해 갈 수 있기 때문이다.

2. 연구방법 및 논문의 구성

본 연구는 주식투자의 방법론에서 시계열 데이터 분석을 통한 행동정책의 방향 설정에 대하여 기존의 고전적 방법론과 최신의 심층학습(Deep Learning)을 이용한 비선형적 분석법을 개괄한다. 이후 최근 연구되고 있는 강화학습 방법 중, Atari Game, 알파고 등 많은 실증 연구가 있는 DQN (Deep Q-learning Network) 방법을 사용하여 데이터 축적이 부족한 경우를 가정하고 GBM(Geometric Brown Movement) Monte Carlo Simulation을 사용하여 가상의 데이터를 통해 학습시킴으로써 실제 충분한 길이의 시계열 데이터 학습효과와 비교하였다. 이를 통해 주식투자 포트폴리오의 강화학습에 가상의 학습데이터 구현의 가능성을 살펴보고자 하였다.

먼저 연구에 사용된 주요 기법들의 개념을 문헌연구를 통하여 정리하고 기본적인 연구비교 구현을 위한 데이터의 정리 방향을 설정하였다. 이를

통하여 분석에 필요한 주식시장 데이터와 종목선정, 그리고 선정된 종목을 통한 포트폴리오의 구성이 어떻게 이뤄지는지 이해하였다. 이후 실제 조사된 방법론을 통하여 10개의 종목을 선정하고 이를 다시 강화학습에 적절한 시계열 길이를 가지고 있는 3개를 선정하여 강화학습에 사용하였다. 이중 강화학습은 준비된 종목 데이터를 사용하여 훈련데이터와 테스트데이터로 나누고 훈련데이터는 비교 분석에 맞춰 데이터 전체를 사용한 훈련과 일부 데이터를 이용하여, GBM Monte Carlo Simulation을 통해 생성된 3개월, 2년의 두 가지 가상의 데이터를 이용하여 모델 훈련하였다. 이후 최근 2개월에 해당하는 선택 3종목 데이터를 추가 수집하여 살펴본 방법론들의 백 테스트를 시행하고 비교하였다.

본 논문은 총 5장으로 구성하였으며, 주요 내용은 다음과 같다.

제1장은 서론으로 본 연구의 배경 및 연구의 필요성을 제시하고 구체적인 연구의 목적과 연구의 방법 및 논문의 구성을 소개하였다.

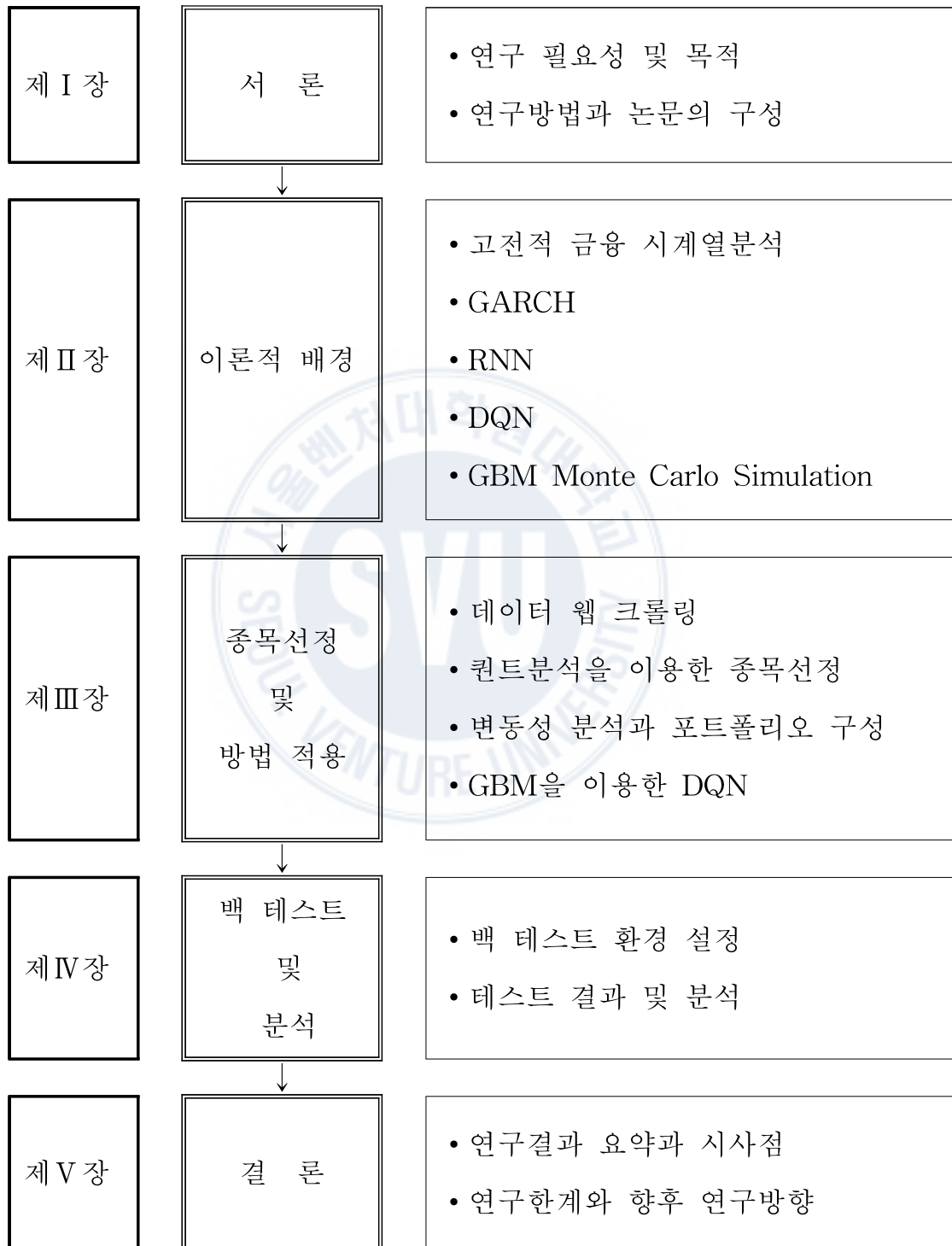
제2장은 연구에 사용된 주요 기법들의 개념들로 고전 계량 투자론의 바탕인 자본 자산 가격 결정모형(CAPM: Capital Asset Pricing Model)과 통계 회귀적 접근 방법, 시계열 데이터 분석법인 지수평활(exponential smoothing)화 방법과 자기회귀 누적 이동평균(ARIMA: Autoregressive Integrated Moving Averag)모형에서 위험률이나 수익률 등의 변동성에 분석 중점을 두고 발전한 일반 자기조건부 이분산성 모델(GARCH: Generalized AutoRegressive Conditional Heteroskedasticity)모형을 살펴보고 다음으로 최근 Deep Learning을 활용한 시계열 또는 연속성 데이터 분석학습 기법인 순환신경망(RNN:Recurrent Neural Network)를 살펴보고 이후 DQN(Deep Q-learning Network)과 GBM Monte Carlo Simulation을 정리하였다.

제3장은 분석에 필요한 데이터를 인터넷을 통한 자료수집 방법인 일명 웹 크롤링(Web Crawling)을 사용하여 대한민국의 주식시장 상장종목의 시계열 주가 데이터 및 종목 대상기업의 재무제표 등의 정보를 수집하고 방법론에 적용할 대상 종목을 선정하였다. 다음으로 2장에서 개괄한 분석 모형 중 CAPM, GARCH, RNN 선택 종목 적용하여 분석하고 포트폴리오를 구성하였다. GBM Monte Carlo Simulation를 이용한 DQN 강화학습 모델에는 포트폴리오 학습 DQN 강화학습 모델을 구현하고 연구를 위한 분석 대상의 데이터를 GBM Monte Carlo Simulation로 2가지 구현, 실제 데이터를 더하여 총 3가지 데이터로 모델을 학습시키고 학습과정 중간 학습결과를 비교하였다.

제4장은 기준일 이후 2개월에 해당하는 다른 변동성의 3종목 데이터를 추가 수집하여 구성한 모델들의 백 테스트(Back Test)를 시행하고 수익률 비교하였다.

제5장은 결론으로 본 연구의 결과를 요약하고, 시사점과 발전적 제언 및 연구 한계점과 향후 연구 방향을 제시하였다.

본 연구의 구성도는 [그림 1-1]과 같다.



[그림 1-1] 연구의 구성도

[그림 1-2] 분석과정 순서도

1) 데이터 수집:

도구	R, Python을 이용
출처	한국거래소 종목정보, 기업정보 사이트의 재무제표 정보, 포털사이트의 주가 정보
대상	코스닥, 코스피 종목
정보	(1차 수집) 기준일자 최신 2년간 데이터 및 강화학습을 위한 1996년부터 수집 가능한 모든 주가 정보



2) 데이터 정리:

작업	코스피, 코스닥 중 우선주, 스펙주, 변동성 없음(거래 없는 주)를 제외
	가치지표에 사용되는 재무비율 계산
	투자지표에 사용되는 가격비율 계산
	모멘텀지표에 사용되는 가격변동성 계산



3) 종목선 선정:

선정	선정 지표 3개(가치지표, 투자지표, 모멘텀지표)에 맞춰 순위 지정
	각 순위에 같은 가중치를 주고 결합 순위 지정
	지정 순위에 따라 10위까지 종목 선정



4) 분석 방법별 포트폴리오 구성 비율 전략 구현:

설정	모든 분석방법을 통한 구현의 수익률은 거래수수료 등의 비용 제외, 비중 계산의 평가는 현금화를 기준으로 계산		
	기준일 과거 2년간의 데이터를 이용 3가지 포트폴리오 구성비율 계산 방법 구현		
분석방법	1) 데이터 통계 분석	2) GARCH 모형	3) RNN을 모형
적용	과거 데이터 기준 기술 통계 과거 데이터를 이용하여 분산 공분산 계산	수집된 2년간의 데이터로 GARCH 모형 만들고 이를 이용하여 기준일 향후 2달간의 종목의 변동성 예측	250일 과거 데이터로 5일 1 step으로 4 step의 20일 예측으로 모형 테스트(이후 백-테스트 모형은 498일 과거 데이터의 10일로 1일 예측 훈련)



[그림 1-2] 분석과정 순서도

5) DQN 구현:

데이터 구분	선정된 종목을 이용하여 GBM MC simulation으로 데이터 생성 구현		
	원본데이터 훈련은 6000일 중 훈련과정을 검증 사용 2000일 데이터를 분리		
	원본데이터 훈련은 훈련 데이터 4000일 검증 2000일 사용		
GBM MC	훈련 데이터 4000일 중 2년간의 데이터를 사용 3개월(63일) window Sliding으로 무작위 선택하고 선택된 window로부터 분산과 초기값으로 향후 2달(63일) 또는 1년(252일) 치 데이터를 생성		
적용 종목 선정	강화학습을 위한 종목 선정 : 가상데이터 학습과 비교를 위해 6000일 이상 길이 및 이후 백 테스트를 위한 변동성 고려 (앞의 10 종목 중 다른 변동 패턴 3개 선정)		
설정	상황(state)은 1일 각 종목의 가격, 보상(reward)은 1일 수익률		
	기초 투자금을 1천만원으로 설정		
	각 종목에 대한 판매, 구매, 유지의 행동(Action)을 통한 구성비 변경		
적용	1) 4000일 원본 데이터 학습	2) GBM MC로 생성된 63일을 이용하여 학습	3) GBM MC로 생성된 252일을 이용하여 학습
	각 데이터 종류 학습에 일정 횟수 학습 시 2000일 데이터를 이용하여 학습 상태 평가		



6) 백 테스트:

데이터 수집	DQN에 사용된 3개 종목(종목길이 및 변동성 고려 선정)의 1차 데이터 기준 일 이후 2달 간의 데이터를 추가 수집 사용		
진행	: 각 방법으로부터 분산, 공분산 계산하고 이를 이용하여 최소변동성을 가지는 포트폴리오 구성비율 계산		
	: 포트폴리오 구성비율 계산 시점의 10일 전의 수익률 정보를 이용하는 것을 기본 전략으로 각 분석방법을 적용		
적용	10일간 각 투자 종목 가격 변동성으로 공분산 계산 후 포트폴리오 구성비 재계산	이전 2년간의 데이터의 GARCH 모형으로 10일마다 예측된 각 종목의 변동성으로 공분산 계산 후 포트폴리오 구성비 재계산	10일마다 각 종목의 앞의 1일 가격변동 예측치와 과거 10일의 가격변동을 합하여 공분산을 계산하고 이로부터 포트폴리오 구성비 재계산
결과	각 방식으로 백-테스트 데이터 2달 기간의 투자 누적 수익률 계산		
	3가지 종류 데이터로 학습된 3가지 DQN 강화학습 모형에 백-테스트 데이터를 넣어 2달 누적수익률을 계산		

II. 이론적 배경

1. 고전적 금융 시계열분석

가. CAPM(Capital Asset Pricing Model)

투자에는 위험과 수익이 발생한다. 투자의 가장 기본적인 명제는 고위험은 고수익을 만들고 반대로 저위험은 저수익을 만든다는 것이다 (Risk-Return Tradeoff: High Risk, High Return). 이를 계량적으로 설명하면 투자하고자 하는 자산의 가치에 변동성의 크기만큼의 위험을 감수한 프리미엄을 더해 수익을 계산하고 실패한 경우 변동성만큼 손해를 보게 된다는 뜻이다. 그리고 이러한 크기를 계량적으로 계산하는 모형이 자본자산가격결정모형(CAPM)이다. 어떠한 자산의 시장가치를 시장참여자들의 투자와 위험의 균형 상태인 자본시장 균형을 이룬다는 가정 아래서 계산하는 모형이다. 본래는 자본시장선(CML:Capital Market Line)과 증권시장선(SML:Security Market Line)을 포함한 개념이지만 주식투자에서는 증권시장선 만을 주로 사용한다.

마코비츠(Harry Max Markowitz: born August 1927-)의 포트폴리오 이론과 샤프(W.sharpe(1964), J.Lintner(1965), J.Mossin(1966)) 등 무위험 자산의 가정을 바탕으로 발전하였다.

$$CML: E(R_p) = R_f + \sigma_p \frac{E(R_M) - R_f}{\sigma_M} \quad (\text{수식 1})$$

$$SML: E(R_i) = R_f + \beta_{im} (E(R_m) - R_f) \quad (\text{수식 2})$$

위의 (식1-1)에 증권시장선(SML:Security Market Line)에서 R_f 는 무위험 이자율을 뜻한다. 자산의 가치가 경제의 성장에 따라 자연스럽게 성장하는 것을 뜻하며 보통 실무에서는 미국 국채 3개월물의 이자율을 사용한다. 일반 상식적으로 금값이 수익률이 이자율이지 않을까 생각하지만 실제로 금값은 일반환율 수익률보다도 변동성이 높아 무위험이라고 할 수 없다. 다음으로 $(E(R_m) - R_f)$ 부분이 시장에서의 프리미엄 즉, 앞에서 설명한 추가적인 위험에 대한 추가수익률 부분이다.

나. 체계적 위험: 베타(β)

이때, β_{im} 은 시장에서 그 자산이 가지는 체계적인 위험을 뜻한다. 이를 수식으로 표현하면 다음과 같다.

$$\beta_i = \frac{cov(R_i, R_m)}{var(R_m)} = \frac{\sigma_{iM}}{\sigma_M^2} \quad (\text{수식 3})$$

주식시장의 변동성분에 주식시장과 개별자산의 공분산으로 이뤄진 베타는 개별자산(주식 종목)의 변동성이 전체 증권시장(주식시장)의 변동성에 대해 얼마나 민감하게 변하는가를 보여준다. SML을 다시 이에 맞춰 해석하면 어떤 자산(주식 종목)의 기대수익률은 무위험 이자율에 추가적인 위험에 대한 프리미엄 가격을 시장에 대한 그 자산의 가치의 변동성의 민감성을 가중해서 더하면 계산 가능하다는 것이다.

CAPM은 주식투자에서뿐만 아니라 유가증권의 가치평가, 자본예산, 투자성과평가 등 재무관리 여러 분야에 걸쳐 광범위하게 사용되고 있다.

이러한 SML의 베타는 회귀분석의 베타와 같은 의미이며, 시장 포트폴리오의 수익률을 독립변수 개별종목의 기대수익률을 종속변수로 놓았을 때 회귀방정식에서의 기하학적 기울기이다.

그런데, 여기서 체계적 위험은 증권시장이 가지고 있는 일반 경제환경의 위험을 뜻하며 분산투자자로 제거될 수 없는 위험을 뜻한다. 하지만 각 종목이 가지고 있는 고유의 기업특성 위험들은 비체계적 위험이라고 하며, 다른 특성을 가진 종목들과의 결합을 통해서 즉, 포트폴리오에 의하여 위험분산이 가능하다.

2. GARCH(Generalized AutoRegressive Conditional Heteroskedasticity)모델

가. 지수평활법과 ARIMA모델

GARCH(Generalized AutoRegressive Conditional Heteroskedasticity)을 이해하기 위하여 지수평활법과 자기회귀누적이동평균(ARIMA: Auto-regressive Integrated Moving Average)모델에 대하여 이해하여야 한다. 고전적인 시계열 데이터의 처리 방법으로 지수평활(exponential smoothing)은 Brown(1959), Holt(1957), Winters(1960) 등에 의해 1950년대 후반 제안되어 시계열 데이터 분석과 예측에 많은 영향을 주었다. 지수 평활은 과거 데이터에 대하여 가중평균(weighted average)하여 미래의 데이터를 예측 또는 평가하는 방법이다. 과거 데이터는 시점이 과거로 멀리 있을수록 지수적으로 감소하는 가중치를 설정한다. 반대로 현시점에 데이터에 높은 가중치를 부여한다. 마치 데이터에 대한 시간적 영향력에 대하여 지수적인 가중치를 설정하는 것과 같다. 이러한 지수평활화 방법은 시간 차이의 영향력을 고려한 데이터 평가의 관점을 정립시켰다고 할 수 있다.

하지만 조심하여야 할 점은 이후 나오는 AR(Autoregressive)에서와 같은 데이터의 스스로에 대한 영향력을 의미하는 것이 아니라 시간적인 가중만을 고려한다는 점이다.

이러한 지수평활 방법은 여러 발전된 방법들이 개발되었는데 가장 기본적인 지수적 가중치를 사용하는 방법을 단순 지수평활(simple exponential smoothing)이라 하며, 식은 다음과 같다.

$$\hat{Z}_{T+1|T} = \alpha Z_T + \alpha(1-\alpha)Z_{T-1} + \alpha(1-\alpha)^2 Z_{T-2} + \dots, \quad (\text{수식 4})$$

여기서 α 는 $0 \leq \alpha \leq 1$ 의 범위를 가지며 평활 매개변수를 뜻한다. 시간 $T+1$ 에 대한 한 단계 앞 데이터(one-step-ahead)는 모든 데이터를 가중평균하여 얻은 값이다. 감소비율은 α 로 조절하게 된다. ARIMA(Auto-regressive integrated Moving Average) 모델은 지수평활 모델과 함께 시계열 데이터에 대한 분석법 중 가장 널리 사용되는 방법이다.

지수평활 모델은 데이터의 추세와 계절성(seasonality)에 대하여 시간적 차이에 대한 영향력을 고려하여 분석하는 반면, ARIMA 모델은 데이터에 시간 또는 연속으로 나타나는 자기 상관(auto-correlation)에 대한 관점을 반영하고 있다.

ARIMA모델을 이해하기 위해서는 정상성(stationarity)과 차분(differencing)에 대한 이해가 필요하다. 정상성(stationarity)은 평균과 분산이 안정적인 상태에서 어떠한 추세나 반복성, 주기 등이 나타나지 않는 일정한 기간이 존재함을 의미한다.

이러한 ARIMA의 정상성은 지수평활 모델과는 다른 점이기에 때문에

두 방법은 서로 보완적으로 사용되어야 한다. 정상성에서 특정한 주기성이나 추세가 보이지 않는다는 것을 랜덤워크(Random Walk)라 한다. 랜덤워크가 보이는 데이터를 정상성을 나타내는 데이터라고 할 수 있다. 이는 관측 시점에 상관없이 각 시점에서 똑같은 분포나 발생 형태를 가짐을 말하고 있다.

자기회귀 모델이 요인 간의 선형 조합을 통하여 인과관계를 설명하듯 ARIMA 모델은 변수의 과거 값의 선형 조합을 이용하여 시계열 데이터를 분석한다. 앞에서 말했듯 지수평활 모델과 다른 점이 바로 이 자기회귀(auto_regressive)를 모델로 한다는 점이다. 시간에 따른 크기의 변화가 아니라 시간의 시점에 따른 자기 자신에 대한 영향력에 대한 분석을 회귀분석으로 접근한다고 할 수 있다.

시간 간격을 차수 p 라 할 때, 이에 ARIMA 모델의 AR 부분 즉, 자기회귀 모델(autoregressive models) 부분은 다음의 식과 같다.

$$Z_t = c + \Phi_1 Z_{t-1} + \Phi_2 Z_{t-2} + \cdots + \Phi_p Z_{t-p} + \epsilon_t, \text{ (수식 5)}$$

이때 ϵ_t 는 백색잡음(white noise)을 뜻한다. p 차 자기회귀 모델인 AR(p)모델이라 한다. ARIMA 모델에서 회귀에서 분석하고자 하는 데이터를 과거 시점의 데이터를 사용하는 대신에, MA(Moving Average)를 이용한 이동 평균 모델을 사용하기도 하는데 이는 앞의 지수평활과 같은 평활화가 아니라 과거 데이터에 대한 모델 자체의 오차(forecast error)을 이용하여 분석하는 모델이다. 이에 대한 식은 다음과 같다.

$$Z_t = c + \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q}, \quad (\text{수식 6})$$

여기에서 ϵ_t 는 백색잡음(white noise)을 나타내고 θ 는 시계열 데이터가 나타내는 계수이다. 이것을 q 차 이동 평균 모델인 MA(q)모델이라 한다.

이렇게 두 가지 모델 AR 모델과 MA 모델을 결합하여 시계열 데이터의 모델을 구성하는 것이 ARMA 모델이다. 여기에 앞에서 설명한 정상성의 확보와 데이터 분석의 용이성을 위하여 추세성과 주기성을 제거하기 위하여 차분(differencing)을 진행하게 되는데 이를 합쳐 ARIMA 모델이라고 한다.

나. GARCH(Generalized AutoRegressive Conditional Heteroskedasticity)모형

앞에서 살펴본 지수평활 모델과 ARIMA 모델은 데이터 자체를 분석하고 예측하는 모델들이다. 그러나 현실에서 사람이 관여된 현상들은 역동적인 영향 관계 변화를 나타낸다. 따라서 데이터 자체의 예측보다 그 변동성을 예측하는 것이 오히려 예측성을 높일 수 있다. 이러한 관점에서 나온 모델이 ARCH(Engle(1982)), GARCH(Bollerslev(1986)) 모형이다.

즉, 주식시장의 종목지수 자체의 변화에 대한 영향요인은 복잡하여 모델에 다 반영하기 힘들지만, 지수의 변동성(volatility)은 일정 기간 큰 영향요인을 받지 않는다는 가정하에 평균과 분산의 변화가 발생하지 않는 변동성 구역(volatility clustering)이 존재하게 된다는 것이다. 이러한 변동성 구역 안에서 변동성 자체에 대한 설명 모델을 구할 수 있다는 것이 조건부 이분산(conditional heteroskedasticity)모형인 ARCH, GARCH 모델이다.

이러한 GARCH 모델은 ϵ_t 는 백색잡음(white noise), ω 를 상수라 할 때 다음의 식과 같다.

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_q \epsilon_{t-q}^2 + \beta_1 \alpha_{t-q}^2 + \cdots + \beta_p \alpha_{t-q}^2$$

(수식 7)

일반적으로 가장 많이 사용되는 모형은 GARCH(1, 1) 모델로

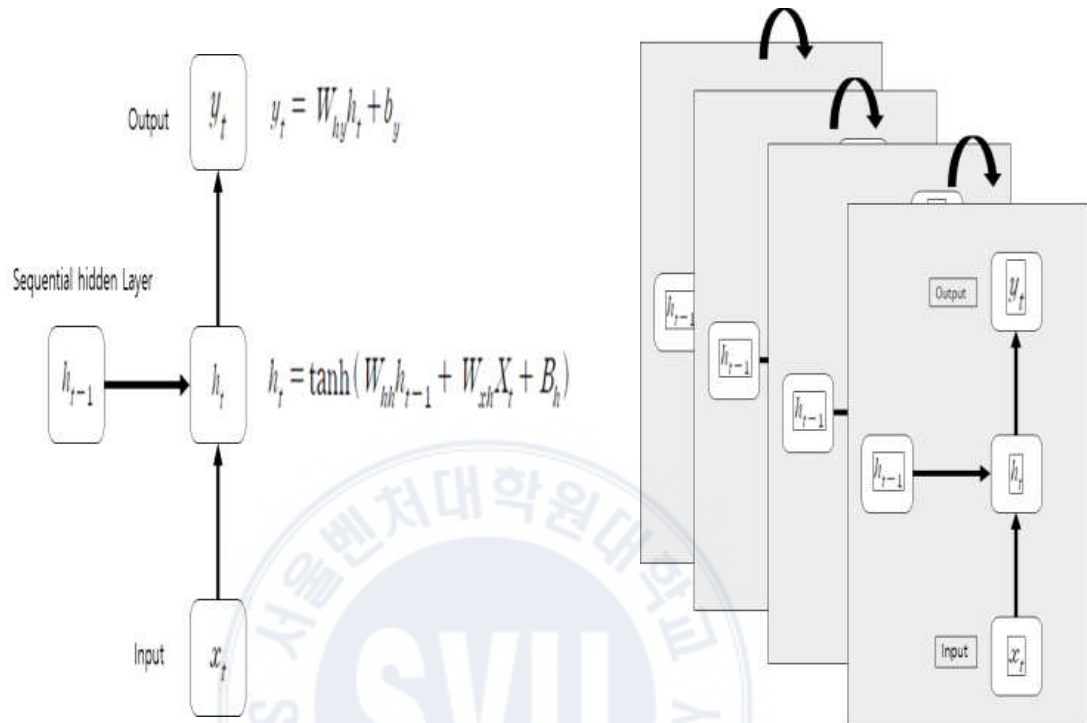
$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \alpha_{t-1}^2 \quad (\text{수식 8}) \text{이다.}$$

GARCH모형은 ARCH모형의 비음(nonnegativity) 조건이 깨질 위험성 등이 있어 이를 보완하기 위해 만들어진 것으로 ARCH(∞) 모형이 GARCH(1,1)와 동치이다. 이후 시계열 데이터분석에서 GARCH를 변형한 GARCH, EGARCH, GJR-GARCH, TGARCH, NGARCH, fGARCH 등이 많이 쓰이고 있다.

3. RNN(Recurrent Neural Network)

가. RNN(Recurrent Neural Network)

순환신경망(RNN: Recurrent Neural Network)은 히든 노드가 방향을 가진 엣지(edge)로 연결돼 방향성 있는 순환을 이루는(directed cycle) 인공신경망(ANN:artificial neural network)의 한 종류이다.



[그림 2-1] RNN의 알고리즘 작동

연속적인 순서가 있는 입력 x 에 대하여 Hidden Layer를 통한 DNN(Deep learning Network)을 형성하고 지난시점의 Hidden Layer의 값을 반복적으로 다시 Hidden Layer의 활성화 값으로 받음으로써 지난 시점의 영향력이 현재의 output 연산에 영향을 줄 수 있도록 설계된 모형이다. 이는 마치 자기회귀(Auto-regressive) 모형을 비선형적으로 구성한 효과를 발휘하고 있다.

이러한 순환구조 특징은 특정한 단위의 순차적 데이터에 대하여 길이와 상관없이 Input과 Output을 받아들일 수 있는 구조를 구성할 수 있다는 것이다. 따라서 유연하게 다양한 구조를 만들 수 있게 되었다. 활

성 함수 (Activate function)로써 하이퍼볼릭 탄젠트(tanh)를 많이 사용하고 있는데 이는 과거로부터의 영향력이 양(+), 음(-) 모두 가능하기 때문이다.

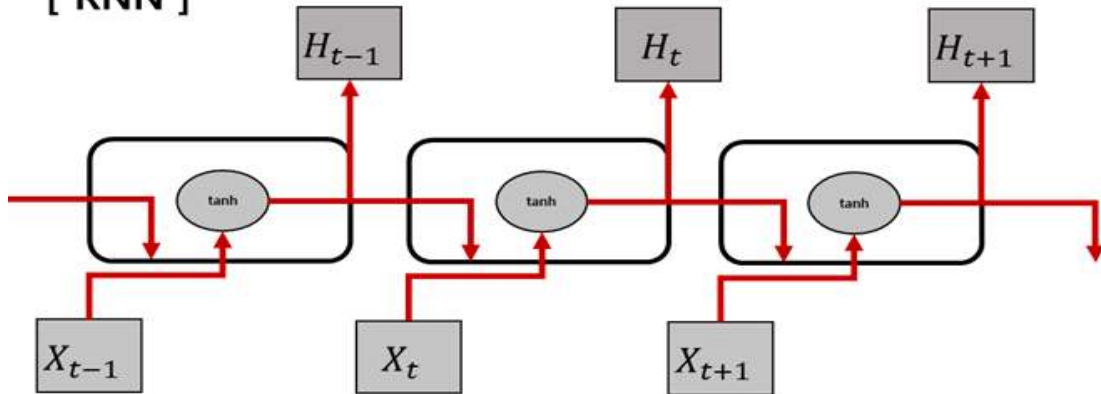
RNN구조는 자연어처리 (NLP: Naural Language Processing)에 주로 많이 연구되어 졌으나 Image 분석이나 Vision 관련 분야에서도 Captioning 등의 분야에서 사용되며, 시계열 데이터 분석 분야에서도 비선형적인 모델을 구성하는데 점점 많은 관심을 받고 있다.

나. LSTM(Long-Shot Term Method)

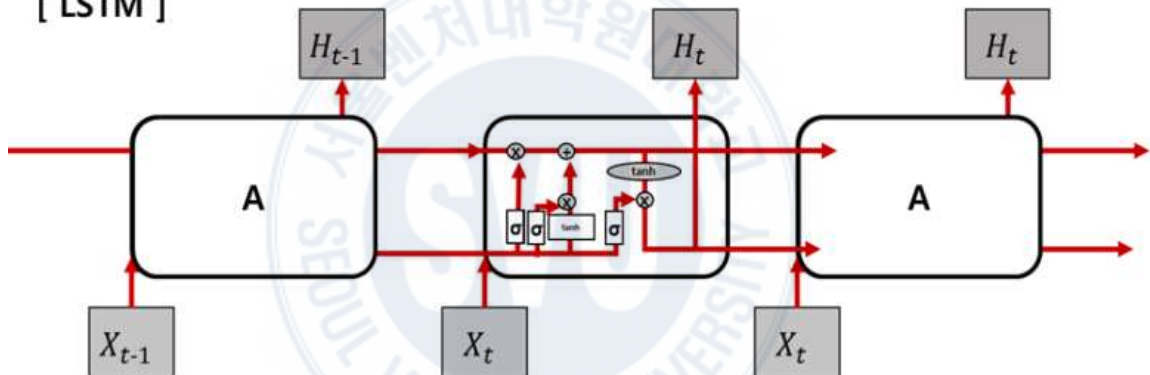
모든 Deep Learning Network가 가진 근본 문제 중 하나는 Network가 깊어지고 학습 과정 중 특정 방향으로의 이상치나 지속적 영향은 학습 Weights의 소멸을 불러오게 된다는 것이다. 이를 vanishing gradient problem이라 한다. RNN은 순환 반복으로 인하여 이러한 vanishing gradient problem이 특히 자주 발생하는 문제가 있다.

그에 대한 대안으로 나온 Network 구조가 LSTM으로 Hidden Layer의 각 State에 cell-state라는 특별한 node를 추가하고 그 안에 forget gate, input gate 등의 과거의 영향력에 대하여 평가할 수 있는 활성화 함수 (Activate function) gate를 통과한 과거의 영향력 정보를 포함하게 하였다. 그에 따라 활성화 함수의 구조는 복잡해 졌지만, 데이터의 길이와 반복성에서 vanishing gradient problem을 해결할 수 있게 됐다.

[RNN]



[LSTM]



[내부 Cell관련 활성화함수]

$$f_t = \sigma(W_{xh_f}x_t + W_{hh_f}h_{t-1} + b_{h_f})$$

$$i_t = \sigma(W_{xh_i}x_t + W_{hh_i}h_{t-1} + b_{h_i})$$

$$o_t = \sigma(W_{xh_o}x_t + W_{hh_o}h_{t-1} + b_{h_o})$$

$$g_t = \tanh(W_{xh_g}x_t + W_{hh_g}h_{t-1} + b_{h_g})$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t$$

$$h_t = o_t \odot \tanh(c_t)$$

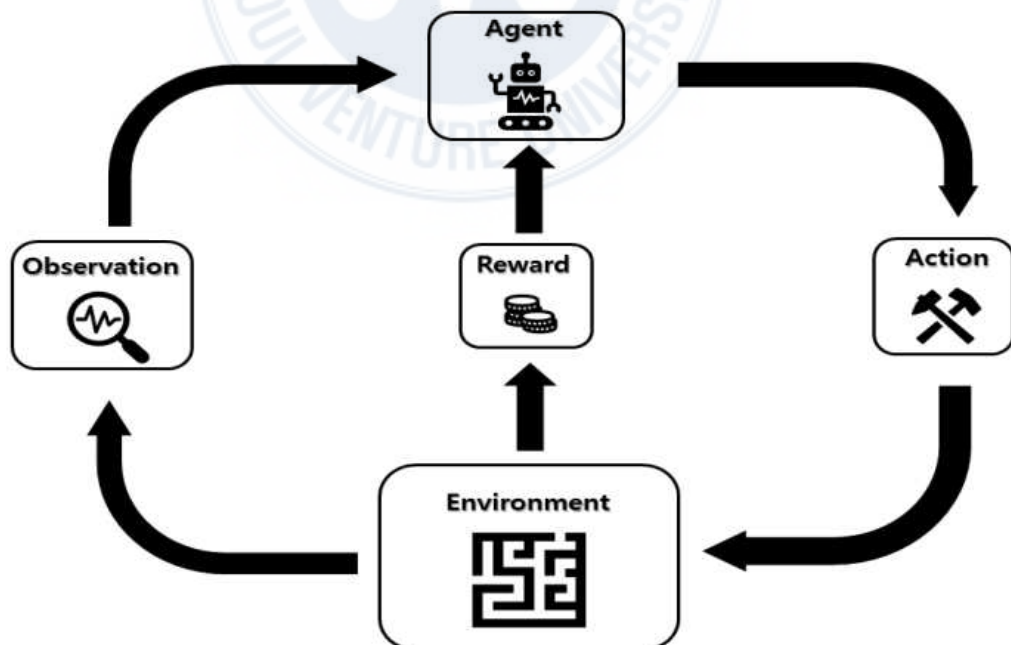
[그림 2-2] RNN과 LSTM 도식과 수식

4. DQN(Deep Q-learning Network)

가. 강화학습 (Reinforcement Learning)

강화학습 (Reinforcement Learning)은 Machine Learning 분야의 한 영역이다. 그 기본 설정은 다음과 같다.

특정 목적을 위한 에이전트(Agent)와 에이전트가 목적을 수행하기 위한 환경(Environment)이 존재할 때, 에이전트는 환경으로부터 특정한 상태(state)에 대한 관측(Observation)을 받아 목적을 향한 행동(Action)을 취하게 되고 그 행동에 대한 상태(state)에 변화와 보상(Reward)을 환경은 새롭게 계산하게 된다. 이를 다시 에이전트는 관측하고 새로운 행동을 취하게 된다.



[그림 2-3] 강화학습 개념도

이러한 설정하에서 에이전트가 목적에 대한 높은 성과를 얻기 위하여, 보상(Reward)을 강화하는 방향으로 행동의 정책(policy)과 상태의 평가(estimate) 방법을 학습해 나아가게 된다는 것이 강화학습이다.

이러한 강화학습은 Markov Decision Process(MDP)에 근간을 두고 확률적인 결과 상황변화에 대한 의사결정과정을 모델화한다. 이는 다섯 가지 측면이 고려되는데 첫째와 두 번째는 앞에서 나온 상태(State)와 행동(Action)이고 세 번째는 한 상태에서 다른 상태로 전이하기 위한 가능성인 상태 전이확률(State transition probability matrix)이다. 이는 다음의 식으로 나타낼 수 있다.

$$P_{aa'}^a = P[S_{t+1} = s' | S_t = s, A_t = a] \quad (\text{수식 9})$$

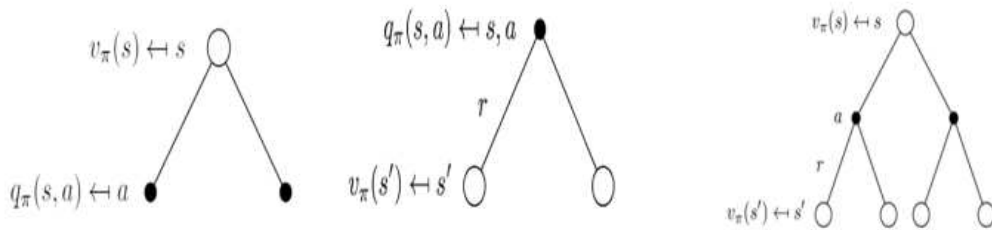
넷째는 보상(reward) 및 각 상태의 보상의 합인 성과(Return) 그리고 성과를 현재 가치화하기 위한 감가율(Discount factor)이다. 이는 다음의 식으로 나타낼 수 있다.

$$R_s^a = E[R_{t+1} | S_t = s, A_t = a] \quad (\text{수식 10})$$

마지막은 에이전트가 행동(Action)을 어떻게 선택할지에 대한 정책(policy)이 있다. 이는 다음의 식과 같다.

$$\Pi(a|s) = P[A_t = a | S_t = s] \quad (\text{수식 11})$$

이러한 관계를 정리한 것이 Bellman Equation으로 다음과 같다.



$$v_{\pi}(s) = \sum_{a \in \mathcal{A}} \pi(a|s) q_{\pi}(s, a) \quad q_{\pi}(s, a) = \mathcal{R}_s^a + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}_{ss'}^a v_{\pi}(s') \quad v_{\pi}(s) = \sum_{a \in \mathcal{A}} \pi(a|s) \left(\mathcal{R}_s^a + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}_{ss'}^a v_{\pi}(s') \right)$$

[그림 2-4] Bellman Equation

상태(state)에 대한 정책의 평가 $V_{\pi}(S)$ 는 각각의 행위(action)의 기댓값에 정책확률(정책에 의해 행위(action)가 선택될 확률)을 곱한 경우의 수 합이며, 행위(action)에 대한 기댓값 ($q_{\pi}(s, a)$)은 행위(action)에 따른 상태 전이확률에 전환된 상태에서 받을 보상(reward)을 곱한 경우의 수의 합해서 현재 시점의 보상 (Reward)을 더해 계산한다.

이들의 관계를 다시 하나의 관계식으로 정리하면 <그림2-4>에서 세 번째 그림과 같다.

나. DQN(Deep Q-learning Network)

앞에서 설명한 MDP는 상태전이확률을 사전에 알아야 하며, 그에 따라 보상이 정해져야 최적 정책을 구할 수 있다. 이러한 환경에 대한 구조와 확률을 알 수 없는 경우 사용되는 방법이 Q-Learning 알고리즘이다. 상태(state) s에서 행동(Action) a를 수행한 결과를 상태(state)들에

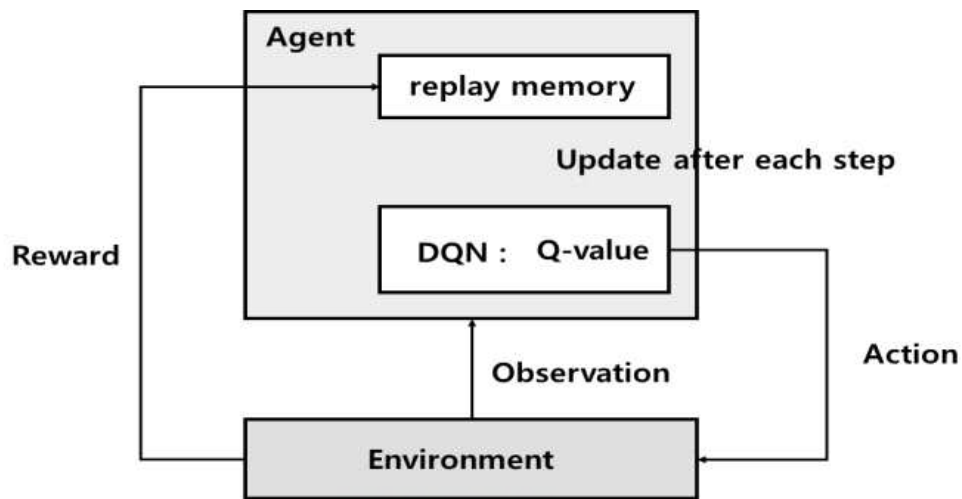
대한 행동(Action)을 제어하는 가치로 사용하는 것이다. 이러한 상태-행동 가치($Q(s,a)$) 큐값이라 하며 이는 다음의 식을 통하여 학습과정을 통해 업데이트 된다.

$$Q(s,a) = Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a') - Q(s,a)) \quad (\text{수식 12})$$

이러한 학습 방법은 식에서 보는 바와 같이 항상 Max한 값으로 이동하려는 경향을 보이기 때문에 탐욕적(Greedy) 정책이라고도 한다. 이러한 탐욕 정책은 새로운 선택에 대한 가능성이 작아진다. 한 번의 학습 결과에 대한 행동 선택은 항상 그 순간의 최고 가치의 행동(Action)을 선택하게 되는 것이다.

이러한 문제를 완화하기 위하여 사용되는 방법이 엡실론-그리드 정책이다. 이는 Q-learning을 일정 부분 탐색적 무작위 행동 선택을 엡실론 비율 만큼 강제적으로 시행하게 함으로써 새로운 행동(Action)에 선택 가능성을 열어 두는 것이다. 그러나 지속적 무작위 행동 선택을 열어 두는 것은 최적 정책을 찾는 것을 불가능하게 한다. 따라서 일정 비율로 엡실론을 감소시켜 학습 진행에 따라 탐색은 줄어들고 탐욕적 정책 선택의 비중을 높이게 된다.

이러한 Q-learning을 바탕으로 구글의 딥마인드사에서 Deep Q-Learning Network(DQN)을 발표하게 되는데 이는 심층학습(Deep Learning Network)에 Q-learning의 큐값을 선형적 알고리즘이 아닌 θ 라는 파라미터를 사용, 비선형적 근사를 통하여 구현하고 그 신경망을 학습시키는 방법이다.



[그림 2-5> DQN 알고리즘

딥마인드사의 DQN은 입력 상태(state)의 관측(Observation)을 CNN을 사용한 Image 분석 Deep Network를 사용한다. 여기서 출력된 tensor 형태의 정보를 Multi-layer 형태의 DQN Agent 모델에 입력시켜 학습하게 된다. 이렇게 산출된 가장 큰 큐값 결과를 Replay Memory라는 저장소에 저장하여 각 Episode 진행 후 무작위로 추출하여 Gradient descent를 계산한다. 이후 특정 batch episode마다 목적 Network를 업데이트시켜 학습하게 된다.

5. GBM(Geometric Brown Movement) Monte Carlo Simulation

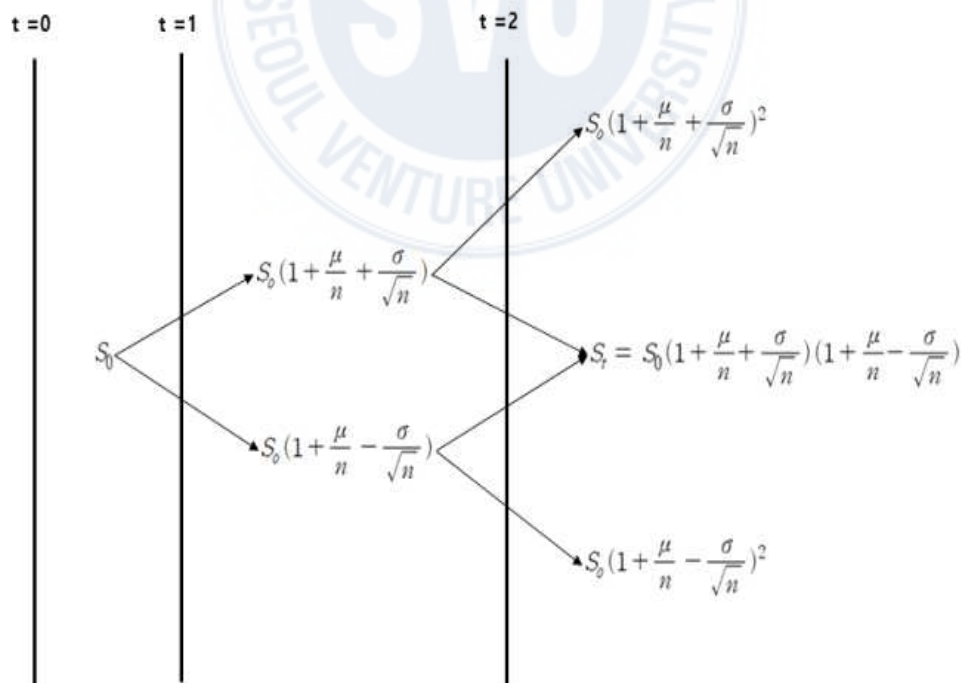
주가의 변화가 일정 시간 간격에서 랜덤하게 움직일 때 이는 마치 액체나 기체 속의 미세입자의 움직임과 같은 무작위 움직임인 브라운운동(Brown Movement: Robert Brown(1827))과 같고 이는 백색잡음(white noise)과도 같다. 이를 바탕으로 무작위 정규분포 난수를 발생시켜 무작

위 변화 데이터를 기존 주가의 변동성과 결합하면 같은 변동성 시간 내에서 무작위 움직임의 주가 simulation을 만들 수 있다. 이를 GBM Monte Carlo Simulation(Paul Samuelson(1950)) 라 한다.

이를 S_0 : 현재 주가로 보고 S_t : t 시점 주가라고 할 때 다음 식으로 나타낸다.

$$S_t = S_0 e^{\sigma W_t + (\mu - \frac{\sigma^2}{2})t} \quad (\text{수식 13})$$

이러한 GBM의 랜덤한 변화를 주가 상승과 하강으로 단순화하여 이해하기 위해 공식을 도식적으로 표현하면 다음과 같다.



[그림 2-6 > GBM의 간략 도식화

이러한 반복적인 랜덤한 방향성을 수식으로 표현하기 위하여

$$S_t = S_0 \left(1 + \frac{\mu}{n} + \frac{\sigma}{\sqrt{n}}\right)^{up} \left(1 + \frac{\mu}{n} - \frac{\sigma}{\sqrt{n}}\right)^{down} \quad (\text{수식 14})$$

(식 13)에 대하여 [총 시행 횟수 : $nt = up + down$], [랜덤한 움직임 : $M_{nt} = up - down$]을 대입하고 다시 $\ln S_t$ 의 로그 형태화하여 근사 정리한 다음 n 의 드리프트 횟수를 무한으로 수렴시키면 처음 (식 12)가 도출된다.

6. RNN, 강화학습 포트폴리오 구성 선행연구

가. RNN 관련 선행연구

RNN을 사용한 연구분야는 자연어처리(NLP: Natural Language Processing) 분야가 가장 중심이 되어 왔다. 자연어처리(NLP)는 인간의 언어나 글을 프로그래밍 언어가 아닌 인간의 언어와 글로 컴퓨터가 바로 입력하고 출력하는 작업을 기반으로 시작하여 인간의 언어와 글을 인공지능이 분석하고 재현하며, 더 나아가 인간의 언어나 글을 매개로 작용하는 작업과정들을 인공지능이 가능하도록 연구하는 분야로 인공지능의 핵심 연구 분야 중 한 분야이다.

RNN은 자연어처리(NLP) 분야를 중심으로 여러 분야에서 시계열 데이터를 분석하거나 예측하는데 적용하려는 연구가 이뤄지고 있다. 예로 감상 후기 또는 강의나 교육의 평가 내용 등의 감성이나 성향을 분석하거

나 기계나 센서의 데이터 해석에 사용되기도 하며, 컴퓨터 프로그래밍에서 프로그래밍 코드 재사용을 이용한 보안 문제 방어 연구, 단조 공정 가열로 가열비용예측, 문서분류 시스템, 각종 제어시스템 연구, 인간의 행동이나 기계의 작동 예측, 환경변화와 기후 변화에 대한 예측 등에 관하여도 적용을 연구하고 있다.

기업경영, 금융투자 분야에서도 제품이나 수요의 예측에 이러한 딥러닝 방법론에 관한 연구가 이뤄졌다. 오병훈(2018)은 딥러닝을 이용한 무기체계 수리 부속의 간헐적 수요예측 연구를 하였다. 박정빈(2019)은 금융시계열 데이터의 수익률 예측에 RNN을 이용하여 금융시계열의 데이터에서 고저가 수익률을 예측하는 연구를 하였다. 김기범(2018)은 RNN을 적용하여 수리적 모델과 딥러닝을 이용한 코스피200 선물변화를 예측하였으며, 조성근(2018)은 암호화폐 가격분석에 암호화폐 분석을 위한 인공지능 기반의 통합모델 제안을 연구하였다, 황승환(2020)은 원/달러 환율예측력 분석 연구에서 전문기관의 원/달러 환율예측력 분석에서 시계열 모형 및 기계학습 LSTM 모형과 비교연구 하였다. 이러한 연구들은 RNN을 이용한 분석과 예측치를 비교하는 연구가 주를 이루고 있다. 본 연구에서는 RNN을 이용한 예측 결과를 포트폴리오라는 행위에 적용하고 그 수익률 결과를 다른 방법과 비교하여 RNN을 통한 투자전략에서의 특성을 파악해 볼 수 있었다.

Machine Learning을 이용한 포트폴리오 구성 종목선정에 관한 연구들도 있었다. 이러한 연구들은 포트폴리오 구성 종목을 선정하기 위한 군집 분류 분석인 클러스터링(Clustering) 분석이 주로 있어 왔다. 박한상(2020)의 연구는 인공지능 기술을 활용한 금융시장 분석과 포트폴리오 최적화에 관한 연구를 하였다. 김경목(2018)은 투자를 위한 특성 집단 분류

를 위하여 머신러닝과 수급분석을 활용한 주식 포트폴리오 구성 연구를 하였다.

이러한 연구들은 기존의 투자대상 지표들을 사용한 기법으로 접근하지 않고 데이터 중심의 포트폴리오 구성 종목선정 방식을 사용하여, 주가나 종목의 정보로부터 주식시장의 종목들의 특성을 분류하여 투자가치가 있는 집단을 구분해내는 방법을 사용하였다. 이들 연구는 본 연구가 기존 투자분석의 지표, 재무분석의 지표 정보를 직접적 순위로 사용하는 것과 달리 순수한 데이터에 기반한 분석으로 종목을 분류하고 선정하는 방식을 취하고 있다.

선행연구들을 살펴보았을 때, 일반적으로 종목선정을 위한 분석과 포트폴리오 구성의 최적화 전략을 연구하는 경우의 두 가지로 연구주제를 분류할 수 있다. 본 연구는 기존 투자지표와 재무지표를 이용하여 포트폴리오 구성에 필요한 종목을 선정하고 이를 이용하여 포트폴리오의 최적화 구성비에 필요한 예측 정보를 RNN 예측을 통해서 구하였다.

나. 강화학습 관련 선행연구 비교

다음으로 강화학습 알고리즘을 이용하여 포트폴리오의 구성 정책(policy)을 자동으로 학습하는 강화학습 선행연구를 살펴보았다. 먼저 강화학습의 정책(policy)과 평가(critic)를 동시에 학습시키는 방법을 사용하는 A3C(Asynchronous Methods for Deep Reinforcement Learning)를 이용한 연구가 있다. 포트폴리오 구성 강화학습에 관한 연구(김주봉, 2019), 비동기 심층 강화학습 기반 금융 포트폴리오 관리와 이를 블록체인 가상화폐에 적용한 연구(김주봉, 허주성, 임현교, 권도형, 한연희, 2019), A3C를 활용한 블록체인 기반 금융 자산 포트폴리오 관리(KIP Trans. Comp.

and Comm. Sys. Vol.8. No.1 pp.17-28 pISSN: 2287-5891)가 있다. 다음으로 본 연구와 같은 DQN 강화학습을 이용한 행위(Action) 중심에 전문가 시스템(Expert System)을 도입한 연구 (Hyung jun Park, Min KyuSim Dong, GuChoi(2019),An intelligentnancial portfolio trading strategy using deep Q-learning, Expert Systems with Applications Volume 158, 15 November 2020, 113573) 가 있다. 다음으로 학습 모델의 행위(Action)에 추가적인 투자전략 행위를 결정하는 전문가 시스템을 도입하고 이를 앙상블 전략으로 수행하는 방법을 연구한 임준범(2019)의 확장된 이산행동 영역에 적용한 강화학습 기반의 행동 특화된 전문가 모델 앙상블 트레이딩 시스템 연구가 있다.

본 연구와 가장 유사한 연구로는 RNN을 이용하여 예측한 데이터를 실제 데이터와 결합하여 입력 데이터로 생성하고 이로부터 포트폴리오 구성비를 정하는 회귀 강화학습 모델에 학습시키는 방법과 가우시안 프로세스(Gaussian Process) 모델을 구성하고 이로부터 큰 방향성은 같으면서 확률적 분포를 가지는 인공 생성 데이터로 원본데이터와 함께 회귀 강화학습 모델을 학습하는 방법을 통해서 성능이 향상되는지 비교 분석한 연구로(강문주, 2018), 회귀 강화학습을 이용한 포트폴리오 자산 배분의 성능 향상 연구가 있다.

본 연구의 같은 점은 첫째, 실측 데이터가 아닌 예측 데이터 또는 가상의 데이터를 생성하여 강화학습에 적용하였다는 점이다. 둘째, 가상의 데이터의 경우 발생의 확률적 분포를 사용하여, 큰 범주의 방향성은 유지하고 확률적 랜덤한 상황의 데이터를 생성하였다는 점이다.

본 연구와의 차이점은 첫째, 본 연구는 확률과정을 가우시안 프로세스를 사용하지 않고 GBM Monte Carlo Simulation 사용하였다는 점이다.

가우시안 프로세스 모델은 적은 데이터로도 일정 범위의 패턴과 예측에 뛰어난 모델로 알려져 있다. 확률적 추론인 베이지안 추론 과정과 같이 모수를 설명하는 함수에 대하여 사전 확률과 사후 확률을 통하여 추정하는 방법을 사용한다. 이 같은 방법을 통하여 불확실성 공간을 정의함이 중요한 분야에서 확률적인 접근에 많이 사용되고 있다. 따라서 가우시안 프로세스 모델로 생성된 데이터는 예측성이 강한 simulation 데이터라 할 수 있다. 반면 본 연구에서 사용한 Geometric Brown movement Monte Carlo Simulation은 방향성 이외에는 랜덤한 정규분포확률을 나타낸다. 즉, 가우시안 프로세스 모델보다 예측성은 떨어지며 동시에 더 무작위적이라 할 수 있다. 둘째, 앞에서 생성한 가상의 simulation 데이터를 원본 데이터와 혼합하지 않고 완전한 가상의 데이터를 이용하여 학습을 진행하였다는 점이다.

이 두 가지 차이점의 목적은 강화학습 알고리즘이 좀 더 가상의 상황에 맞는 학습을 진행하고자 하였기 때문이다. 즉, 현재의 데이터로부터 가능성 있는 확률이나 원본데이터로부터의 정보가 배제된 가상의 데이터로 학습하기 위함이다. 이는 본 연구의 목적이 가상의 환경 조건 설정으로부터 만들어낸 simulation 데이터로도 강화학습이 학습 가능한지를 확인하고자 하기 때문이다.

반면, 학습의 성능을 테스트하기 위하여 현실성 있는 변동성이 심한 최근 국내 주가 데이터를 사용하여 학습상태를 검증하고자 하였다. 다른 패턴의 움직임을 보이는 3개 종목을 선정하여 백 테스트를 진행하고 가상 데이터의 학습결과인 누적수익률의 발생 결과를 확인하였다.

Ⅲ. 종목선정 및 방법 적용

1. 데이터 웹 크롤링(Data Web Crawling)

가. 데이터 웹 크롤링(Data Web Crawling)

본 연구는 실험 대상 데이터로 코스피(KOSPI), 코스닥(KOSDAQ) 주가 데이터를 기준으로 분석하였다. 한국거래소 정보와 포털사이트에서 제공하는 금융정보, 기업정보제공 회사의 재무제표 정보를 웹 크롤링 기법을 이용하여 R 프로그램 및 Python 프로그램으로 수집하였다. 주가는 수정주가(배당을 반영한 가격)가 아닌 종가(Close) 사용하였고 주식구매를 통한 인수 합병을 목적으로 공모를 통해 투자자로부터 조달한 자금으로 만든 회사 종목(SPAC: Special Purpose Acquisition Company)과 의결권 없이 배당 및 이익 배분을 목적으로 하는 우선주 종목은 제외하였다.

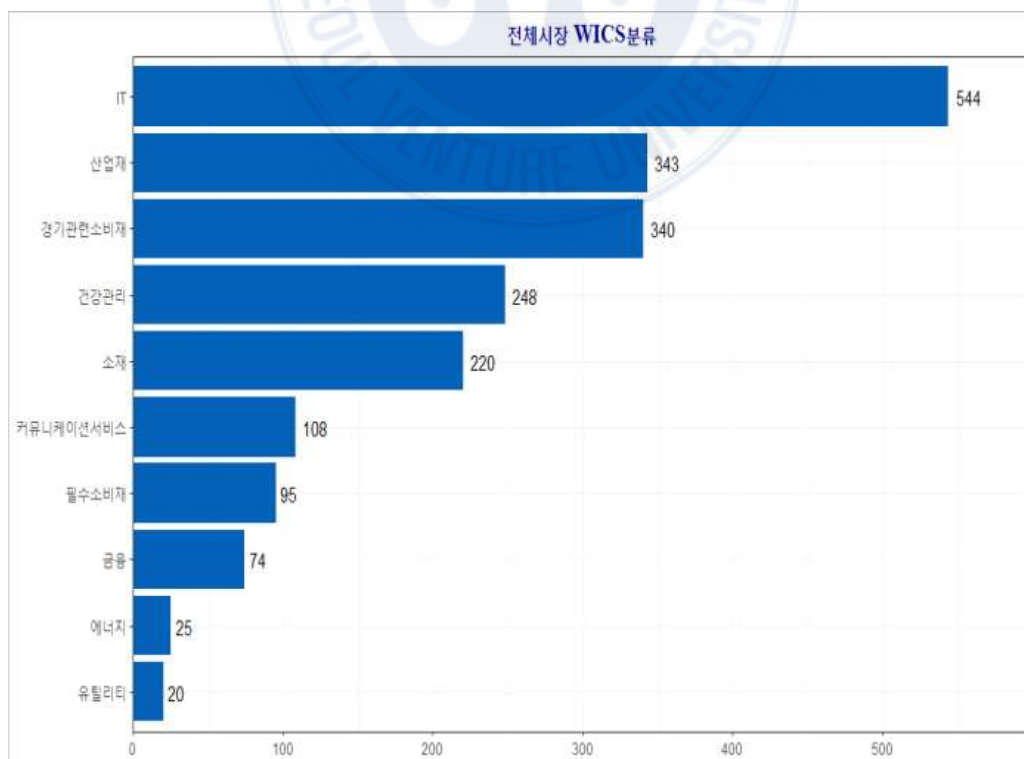
한국거래소 웹사이트에서 ‘마켓데이터조회’를 이용 산업별 현황 데이터를 내려 받아 분석에 필요한 증권시장의 종목분류와 종목별 각종 지수 및 발행주식 수 등을 확인하여 활용하였다. 한국거래소의 정보를 바탕으로 국내 포털사이트에서 제공하는 금융 및 증시에 관련한 데이터에서 분석 기간에 대한 대상 주식 가격정보를 R 및 Python 크롤링 기법을 이용하여 확보하였다. 기업투자 정보제공회사 웹사이트에서 제공하는 기업별 재무제표 중 포괄손익계산서(연간), 재무상태표(연간), 현금흐름표(연간)를 확보하였다. 이렇게 확보한 재무자료로부터 데이터를 정리하여 가치지표인 PER, PBR, PCR, PSR 등을 산출하였다.

주가 데이터의 기준일자는 2020년 7월 21일이고 재무제표는 2019년 12

월 기준일로 수집하여 모델에 사용할 종목선정과 모델 구성에 사용하였다. RNN과 DQN구현에 사용된 데이터는 해당 데이터에 대하여 별도로 Sliding window 방법을 적용하여 학습에 사용하였다.

나. 데이터 정비(Data Cleaning and Prepare)와 이해

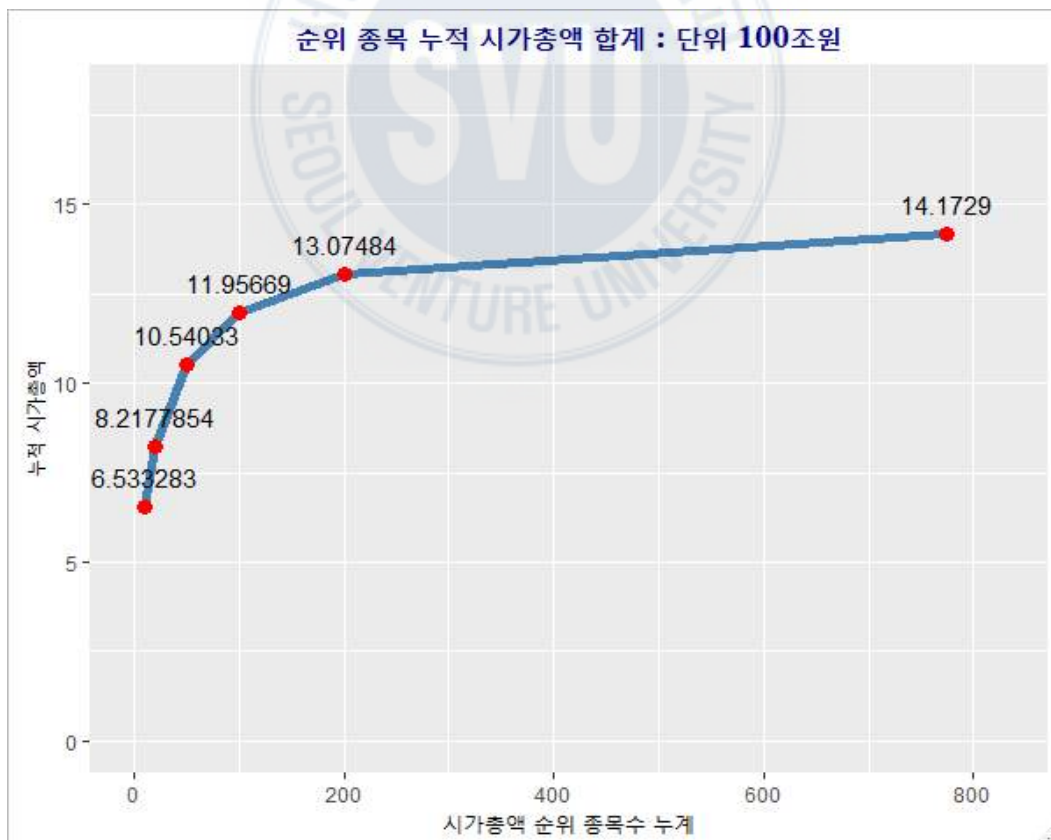
데이터의 이상치를 제거하는 방법에는 이상치 데이터를 범위 기준치나 평균값으로 대체하는 윈저라이징(Winsorizing) 방법과 이상치 범위 한계를 넘은 데이터를 제거하는 트림(Trim) 방법이 있다. 본 연구에서는 이상치 제거가 필요할 경우 윈저라이징을 사용하였다. 가치지표를 비중으로 사용해야 하는 경우는 왜곡을 막기 위해 표준(z-score) 정규화를 사용하였다.



[그림 3-1] 전체시장 WICS 분류

수집된 주식 종목 수는 총 2116개로 이중 KOSPI는 775개 종목이며 나머지는 KOSDAQ 종목이다. [그림 3-2]는 주식시장 전체 대상 종목을 산업별로 분류한 것이다.

주식의 산업을 분류할 때 기준은 MSCI와 S&P가 개발한 ‘GICS12’가 가장 많이 사용되나, 독점적 지적재산으로 명시되어 있어 저작권 허락 없이 사용할 수 없다. 따라서 본 연구에서는 한국거래소 산업 현황 자료를 사용하여 ‘WICS’ 기준으로 산업을 분류한다. ‘WICS’는 지수 제공업체 ‘와이즈인덱스13’에서 제공 발표하는 GICS와 비슷한 분류기준이다.



[그림 3-2] 순위 종목 누적 시가총액 합계

총 KOSPI 주식시장 시가총액은 약 1,420조원(스팩, 우선주 제외) 상위 10개 기업의 전체 주식시장 시가총액 중 약 650조원(약46%)을 차지하며, 상위 20개 기업까지의 전체 주식시장 시가총액은 820조원(약58%) 정도이다. 상위 200대 기업의 시가총액이 전체 주식시장에서 차지하는 금액은 약 1,050조원(약74%)이다. KOSPI200 지수의 선정이 기준이 시가총액만은 아니다. 그러나 전체 종목의 비중으로 보았을 때 200대 기업이 전체시장을 대표한다고 볼 수 있다.

10대 기업에서 20대 기업이 국내 주식시장에 50%를 차지한다는 것은 산업의 안정성과 다양성에 비춰볼 때 국내 기업 환경과 경제에 대하여 시사하는 바가 있다.

2. 켄트분석을 이용한 종목선정

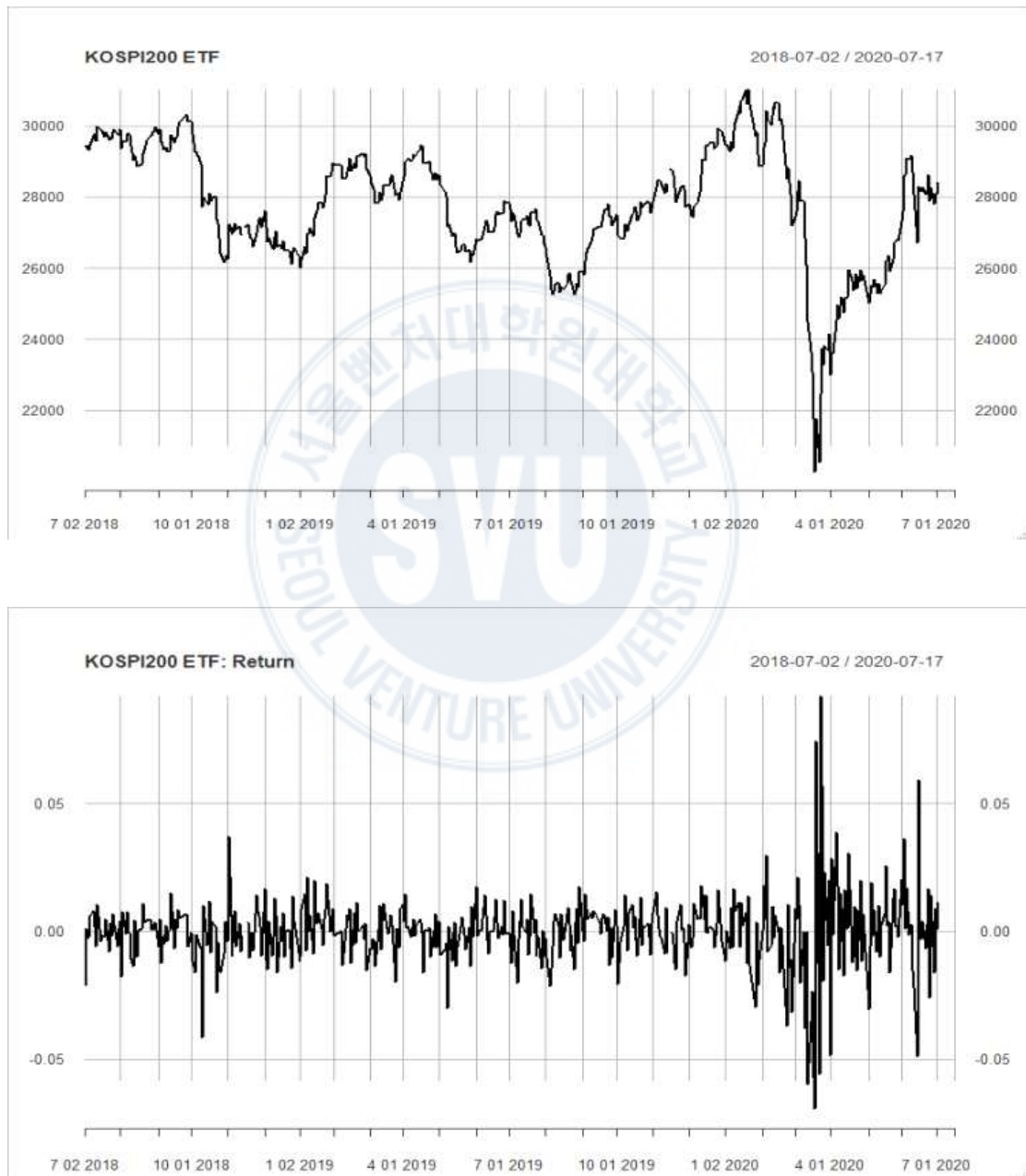
가. 종목선정 기준

a. 베타와 변동성 기준

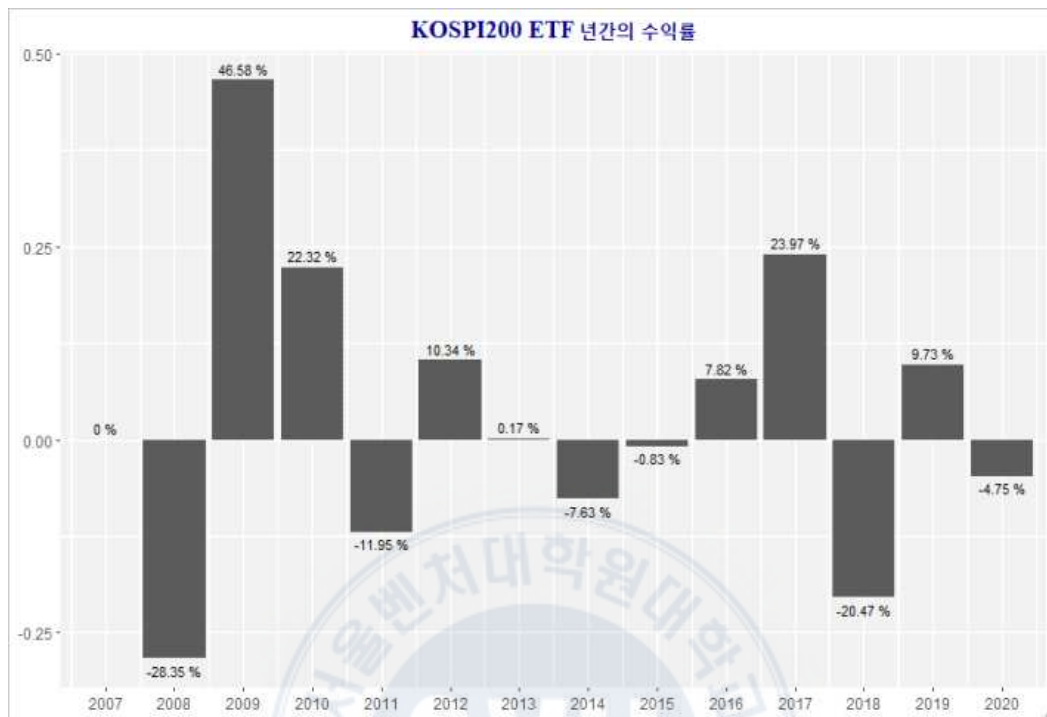
주식시장의 베타(β)에 관하여 KOSPI 및 KOSDAQ에 직접적용 분석하기는 힘들다. 이를 대신하기 위하여 직접적인 공분산을 사용하지 않고 주식시장 지수 펀드 종목인 KOSPI 200 ETF(Exchange Traded Fund : 상장지수펀드)를 주식시장 시장분석의 비교 기준대상으로 놓고 상대적인 변동성 기울기를 파악하여 종목별 베타(β)를 관찰하였다.

이에 지수 펀드 KOSPI 200 ETF인 TIGER 200(102110.KS)의 2년간 가격 정보를 받아서 사용하였다. 베타(β)를 시장 종목과 비교하기에 앞서

ETF를 이용하여 지난 2년간의 KOSPI의 가격변동과 연간 수익률을 그래프로 만들고 살펴보았다.



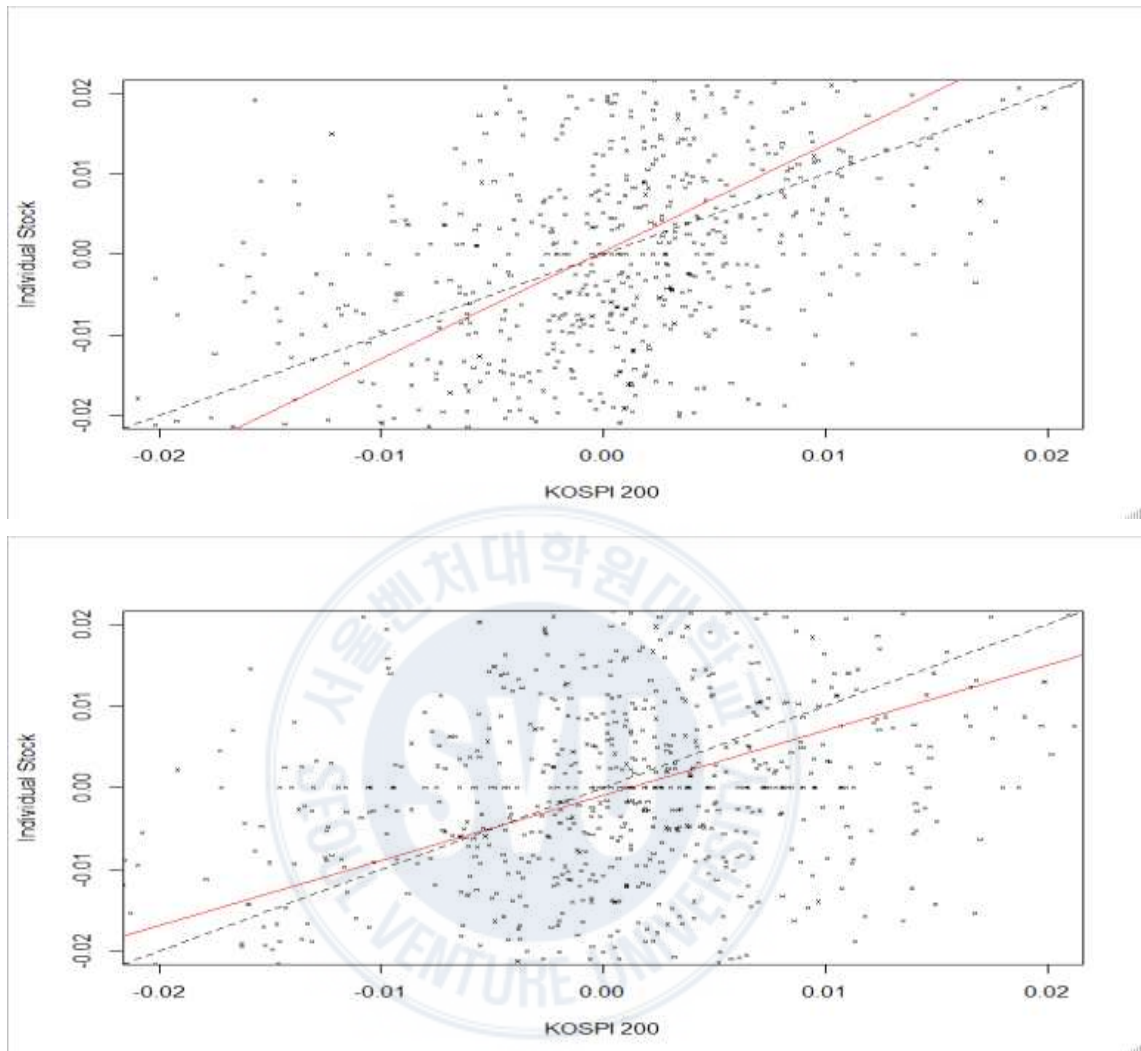
[그림 3-3] KOSPI 200 ETF 2년간 가격변동과 연간 수익률



[그림 3-3] KOSPI 200 ETF 2년간 가격변동과 연간 수익률

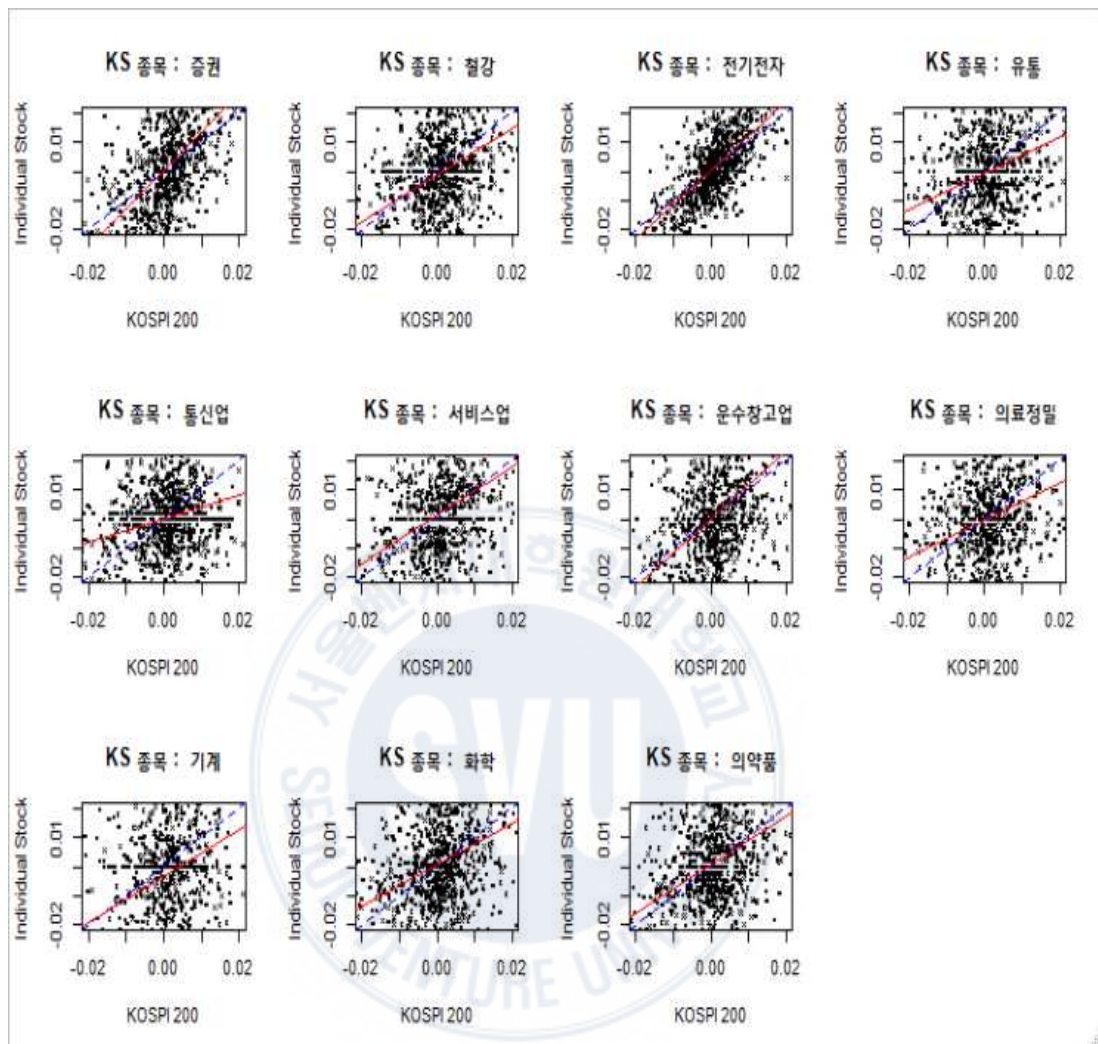
[그림 3-3]의 위에 그래프는 지난 2년간의 KOSPI200 ETF의 가격을 보여주고 밑에 그래프는 같은 기간의 수익률 즉, 변동성을 보여주고 있다. 마지막 그림표는 2008년 이후 매년 년 단위 수익률을 보여 준다.

본 연구에서 종목선정에 베타(β)는 가장 기초적인 회귀분석 모델을 통해서 특정 산업군 또는 종목이 가지는 시장과의 특성을 파악하는 지표로 분석하였다. 베타(β)의 산업별 특성을 살펴보기 위해 앞에서 다룬 KOSPI200 ETF와 키움 증권사 및 포스코 강판의 2017년 1월부터 2020년 7월까지의 주가의 관계를 scatter plot으로 나타내고 그 위에 회귀계수 선을 그어 비교해 보았다.



[그림 3-4] KOSPI 200 ETF와 키움증권 및 포스코강판 베타(β)선

키움증권의 변동성은 주식시장의 변동성보다 크게 나타나고 있다. 반면, 포스코 강판은 주식시장의 수익률 변동성 보다 지난 2년간 변화율이 적다는 것을 알 수 있다. 각 산업 종목의 기울기를 통하여 특정 산업 종목이 시장의 변동성에 비하여 얼마나 크게 변하는지 알 수 있다. 다음은 각 산업군에서 임의로 선택한 기업의 Beta를 표시하였다. 직선은 beta를 점선(대각선)은 코스피 수익률 기준선이다.

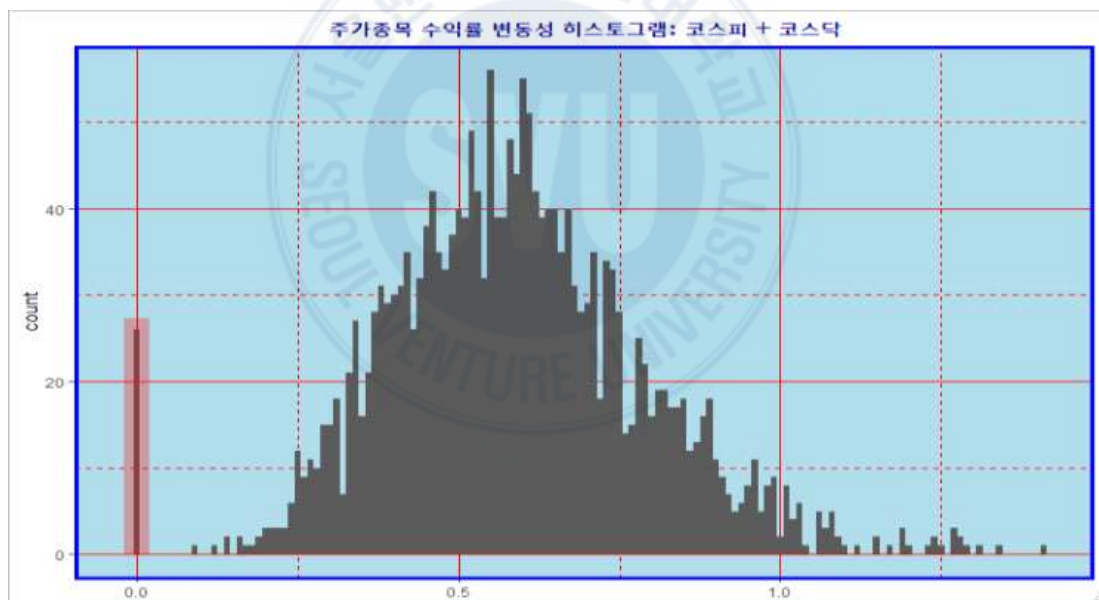


[그림 3-5> 주식시장 산업군 임의 선정 기업 Beta(β)선

이러한 변동성 계수 베타(β)에 대하여 일반적으로 높은 위험, 높은 수익이 투자론에서는 알려져 있다. 즉, 고변동성 종목의 기대수익률이 크고, 저변동성 종목의 기대수익률이 낮다는 것이다.

하지만 현실에서는 저변동성의 종목들에 수익률이 높은 경우도 많이 존재하며 이러한 현상을 브르네미어와 베이커(Brunnermeier and Parker 2005)는 투자자들의 자산에 정보를 과신하는 경향성이 고변동성을 과대평

가하게 되어 결국 수익률을 낮추는 결과를 가져온다고 보았으며, 베이커와 브레들리, 버글러(Baker, Bradley, and Wurgler 2011)는 기관투자자들이 레버리지 투자가 되지 않는 상황에 벤치마크 대비 높은 성과를 얻기 위해 고변동성 주식에 투자하는 경향이 고변동성 주식을 과대하게 평가해 결국 수익률이 낮게 나오는 결과를 만든다고 설명하고 있다. 반면 세프톤(Sefton et al. 2011)은 시장은 상승과 하락이 반복되고 고변동성 주식에 변동성 손실(Volatility Drag)을 지속적으로 받게 되며 이로 인해 수익률이 하락하게 된다고 보았다.

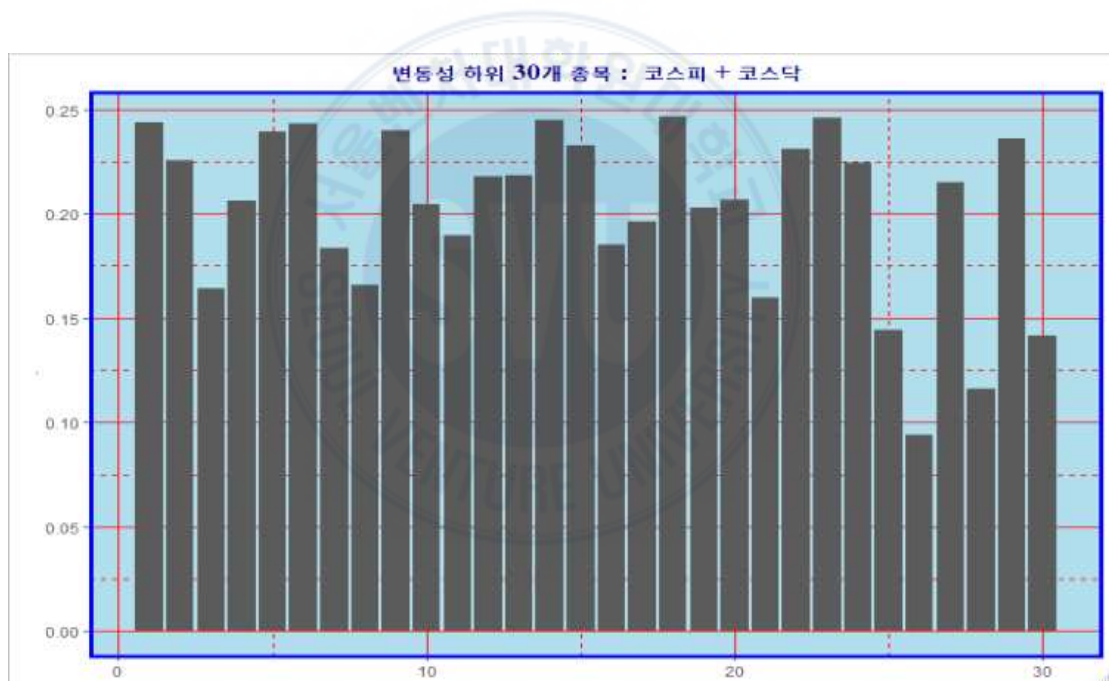


[그림 3-6] 주가종목 수익률 변동성 히스토그램

따라서 본 연구의 종목 선택에 저변동성을 기준으로 삼고자 한다. [그림 3-6]은 최근 1년 일간 수익률 기준 변동성이 낮은 주가 종목을 확인하기 위하여 데이터에서 영업일 기준 252일(평균 연간 시장 업무일 수)의 주가 수익률하고 거기에 표준편차를 구하였다. 표준편차는 분산의 양의 제곱근

을 뜻하며 이것은 변동성의 크기를 나타내게 된다. 이를 히스토그램으로 나타낸 것이다.

[그림 3-6] 주가종목 수익률 변동성 히스토그램에서 0에 위치하는 종목들은 최근 1년간 거래정지 등으로 인해 가격이 변하지 않았고, 이로 인해 변동성이 없는 종목들로 해당 종목들은 변동성 분석에서는 제외하였다. 이후 변동성이 적은 30개 종목에 대하여 히스토그램으로 변동성을 살펴보면 [그림 3-7] 주가종목 수익률 변동성 히스토그램과 같다.



[그림 3-7] 수익률 변동성 하위 30위 히스토그램

b. 모멘텀(Momentum) 기준

다음 종목선정의 기준으로 모멘텀(Momentum)을 사용하였다. 투자에서 모멘텀이란 주가 혹은 수익률 추세를 말한다. 다시 말해 변동성의 추세로서, 상승 추세의 주식은 일정 시간 동안 지속 상승하며 하락 추세의 주식

은 일정 기간 지속 하락하는 현상을 말한다. 모멘텀(Momentum) 현상에 대한 가장 일반적인 설명은 투자자들의 스스로에 대한 과잉 신뢰 또는 특정 믿음에 대한 고착이나 관성 때문으로 본다.

투자자들은 자신의 판단을 근거가 되는 정보를 과하게 평가하고, 자신의 판단을 부정하는 정보에 대해서는 과소하여 평가하는 경향이 있다는 것이다. 이러한 심리적인 요인이 모멘텀 현상을 발생시킨다고 보고 있다.

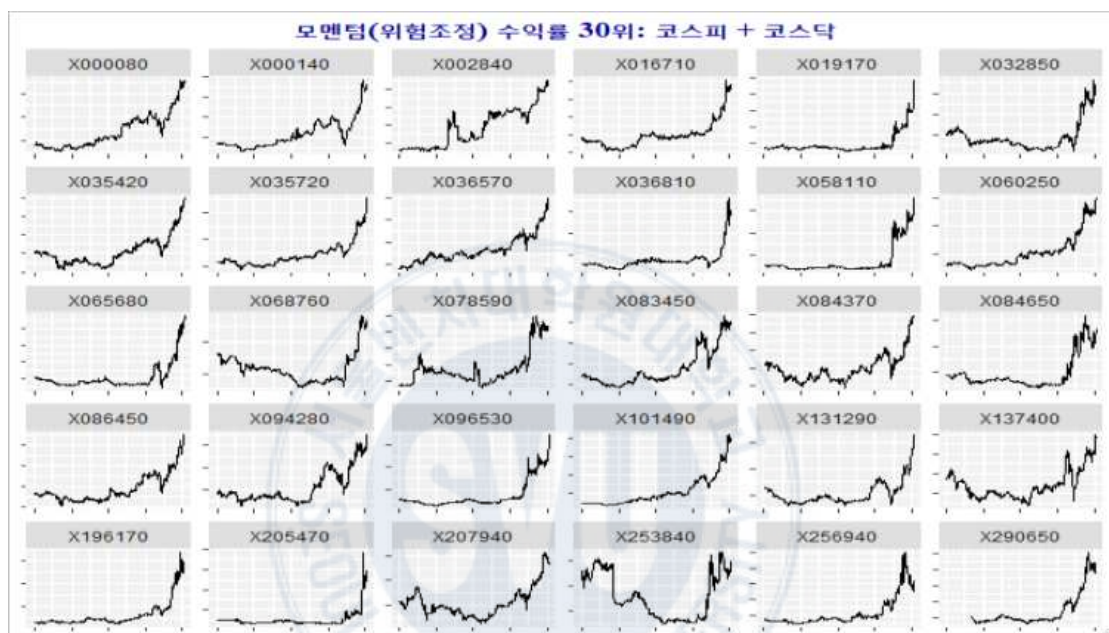
모멘텀에는 수익률 즉, 변동성의 추세에 관한 모멘텀과 주가 자체의 모멘텀으로 나뉘며 주로 모멘텀이라고 이야기되는 것은 3개월에서 12개월을 의미하는 중기 모멘텀(Jegadeesh and Titman 1993)의 주가 모멘텀을 이야기한다. 변동성 추세를 나타내는 수익 모멘텀(Rendleman Jr, Jones, and Latane 1982)과 장기 3년에서 5년을 의미하는 모멘텀(De Bondt and Thaler 1985) 단기 1주일(Lehmann 1990) 혹은 1개월 이하(Jegadeesh 1990) 모멘텀에 대한 연구도 있다.

그런데 이러한 가격 모멘텀은 안정적이지 않으며 지속해서 내외부의 이벤트나 영향요인으로 급변동을 일으키기 쉬우며 이는 데이터를 왜곡시킬 수 있다. 따라서 일정 기간의 누적수익률에 변동성(분산 또는 표준편차)을 나눠 수익률의 추세나 가격의 추세에 위험을 고려할 수 있다. 이러한 방법을 위험(변동성)조정(Risk-adjustment)이라 한다. 이러한 위험률 조정 후의 모멘텀은 샤프지수(William Sharpe, 1966)와 같은 위험조정수익률 계산과 유사해진다. 이해를 위해 아래는 사전 샤프지수의 표현식이다.

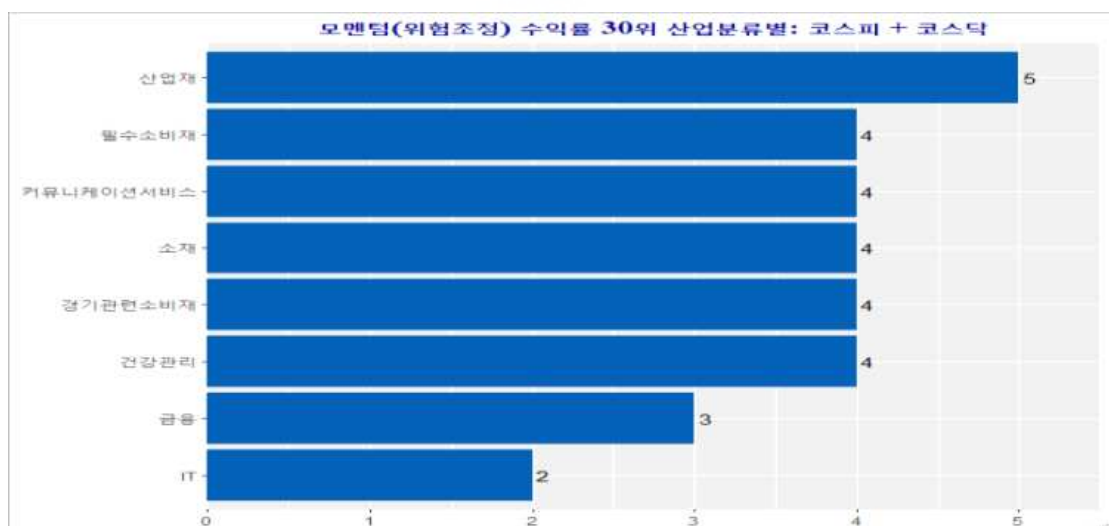
$$sharp비율 = (기대수익률 - 벤치마크수익률)/표준편차 \quad (수식 15)$$

다음의 [그림 3-8]은 지난 12개월의 모멘텀을 계산하고 위험조정 하였

을 때, 모멘텀이 높은 순위 30위의 주가를 그래프로 나타낸 것이다. <그림 12>에서 보는 바와 같이 대부분 안정적인 우 상향 형태를 보임을 확인할 수 있다. [그림 3-9]는 산업별로 위험조정 모멘텀 30위를 분류한 것이다.



[그림 3-8] 모멘텀(위험조정) 수익률 30위 주가 그래프



[그림 3-9] 모멘텀(위험조정) 수익률 30위 산업분류별

c. 투자가치 기준

주가와 발행주식 수 그리고 재무제표로부터 투자를 위한 여러 지표가 개발되어 있다. 이러한 지표들을 통해서 종목을 선정하는 기준으로 삼을 수 있다. 특히 많이 주목받는 투자지표들은 PER, PBR 등으로 이러한 지표는 기업의 내재 가치를 나타내고 있어, 내재 가치 대비 낮은 가격의 주식을 찾아 투자할 수 있기 때문이다. 이러한 효과를 가치주효과(Basu 1977)라 하며 내재 가치 대비 낮은 가격의 주식에 투자하는 것이 내재 가치 대비 비싼 주식보다 수익률이 높은 효과를 말한다. 반대로 성장주란 현재의 가치는 낮지만 성장가능성을 높게 보아 가격이 높게 평가되는 주식을 말한다.

그러나 이러한 지표는 명확한 표준 기준이 있지 않다. 투자자나 분석가의 기준이 다를 수 있고 또 시장이나 산업분류의 특성을 고려하여야 한다. 예를 들어 PER, PBR, 등의 지표는 기업소유를 위한 지주사가 그 특성상 많이 포함되게 된다. 다음은 투자가치 지표의 계산식이다.

PBR(Price to Book-value Ratio:주가순자산비율) = 주가 / 주당 순자산 (수식 16)

PER(Price to Earnings Ratio:주가순이익비율) = 주가 / 주당 순이익 (수식 17)

PCR(Price to Cashflow Ratio:주가현금흐름비율) = 주가 / 주당 현금흐름

= 주가 / 영업현금흐름(OCF: Operating Cash Flow)

= 주가 / 잉여현금흐름(FCF: Free Cash Flow) (수식 18)

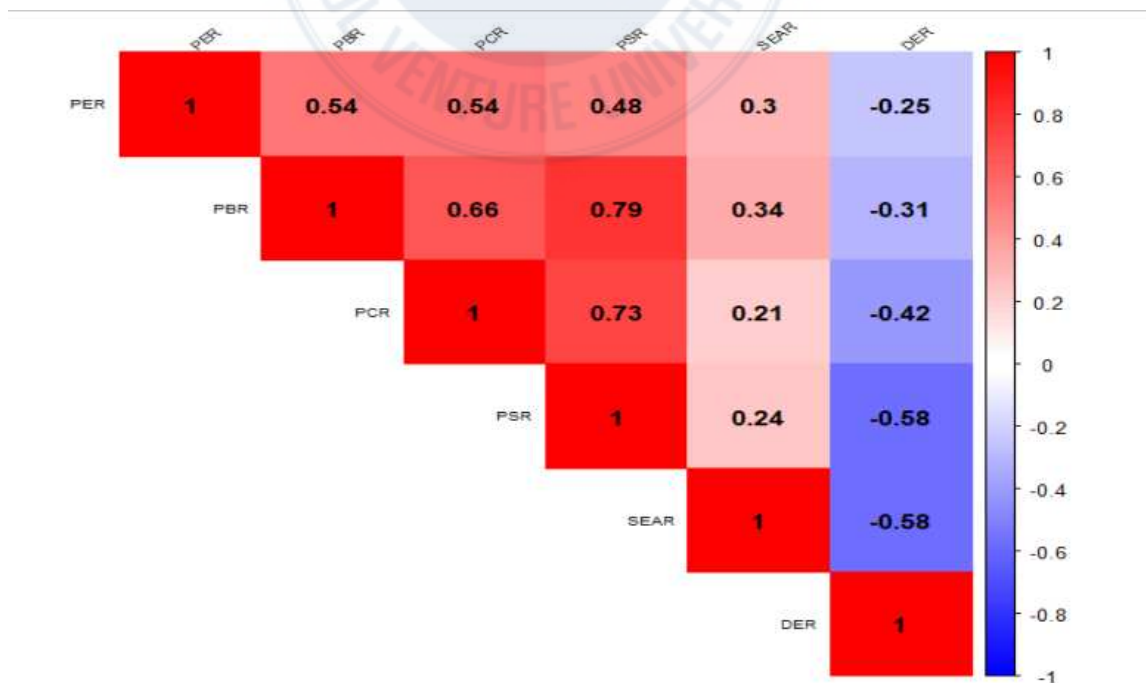
PSR(Price to Sales Ratio: 주가매출비율) = 주가 / 주당 매출액 (수식 19)

본 연구에서는 수집한 재무제표의 항목과 주가 데이터에서 PER, PBR, PCR, PSR을 구하고 각 종목에 각 지표에 대한 등수를 부여하여 순위를

구하였다. 또한, 지표 결과를 바탕으로 지표 간의 상관계수를 구하였다. 그래프의 SEAR(자기자본비율: Stakeholder's Equity of Total Asset Ratio) 와 DER(부채비율: Debt to Equity Ratio)을 추가하여 비교하기 쉽게 하였다.

저평가 주식이 계속해서 저평가에 머무르는 다른 이유가 존재하기 때문에 이러한 가치 함정에 빠지지 않으려면 다른 관점의 여러 지표를 동시에 볼 필요가 있다.

[그림 3-10]에서 볼 수 있듯이 일반적으로 종목을 선택하는데 가치지표들을 구해서 데이터의 내용을 분석할 수 있다. 이들 지표는 일반적으로 많이 사용되는 종목 선택의 지표들이지만 이들 간에도 상관계수가 0.5 이하로 같은 양의 관계이더라도 그 관계에 강도가 다르다는 것을 볼 수 있다.



[그림 3-10] 종목선정 투자가치 지표 간의 상관계수

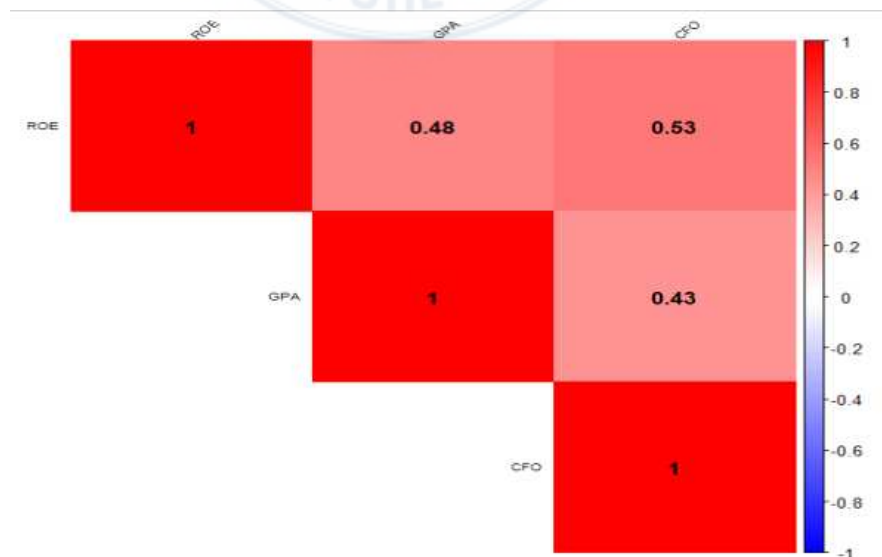
d. 재무지표 기준

기업 재무 건전성 지표라고도 한다. 재무제표 분석을 통한 기업 자체의 재무적인 정보를 이용하여 투자의 기준으로 삼는 것이다. 이러한 재무제표 정보에 대한 지표는 이미 회계나 재무관리 등에 의해서 오랜 시간 공인되어 온 지표들이다. 이러한 지표들은 투자 분석가나 회계, 재무전문가에 의하여 전문적이거나 새로운 지표를 종목과 산업 분야에 맞게 고안할 수도 있다. 본 연구에서는 가장 일반적인 자기자본이익률(ROE), 매출총이익(Gross Profit), 영업활동현금흐름(Cash Flow From Operating)을 중점으로 주식시장을 분석하고 종목 선택의 기준으로 삼았다. 다음은 사용된 재무제표 지표 계산이다.

$$\text{ROE}(\text{Return On Equity: 자기자본이익률}) = \text{지배주주순이익} / \text{자본} \quad (\text{수식 20})$$

$$\text{GP/A} = \text{매출총이익} / \text{자산} \quad (\text{수식 21})$$

$$\text{CFA} = \text{영업활동으로 인한 현금흐름} / \text{자산} \quad (\text{수식 22})$$

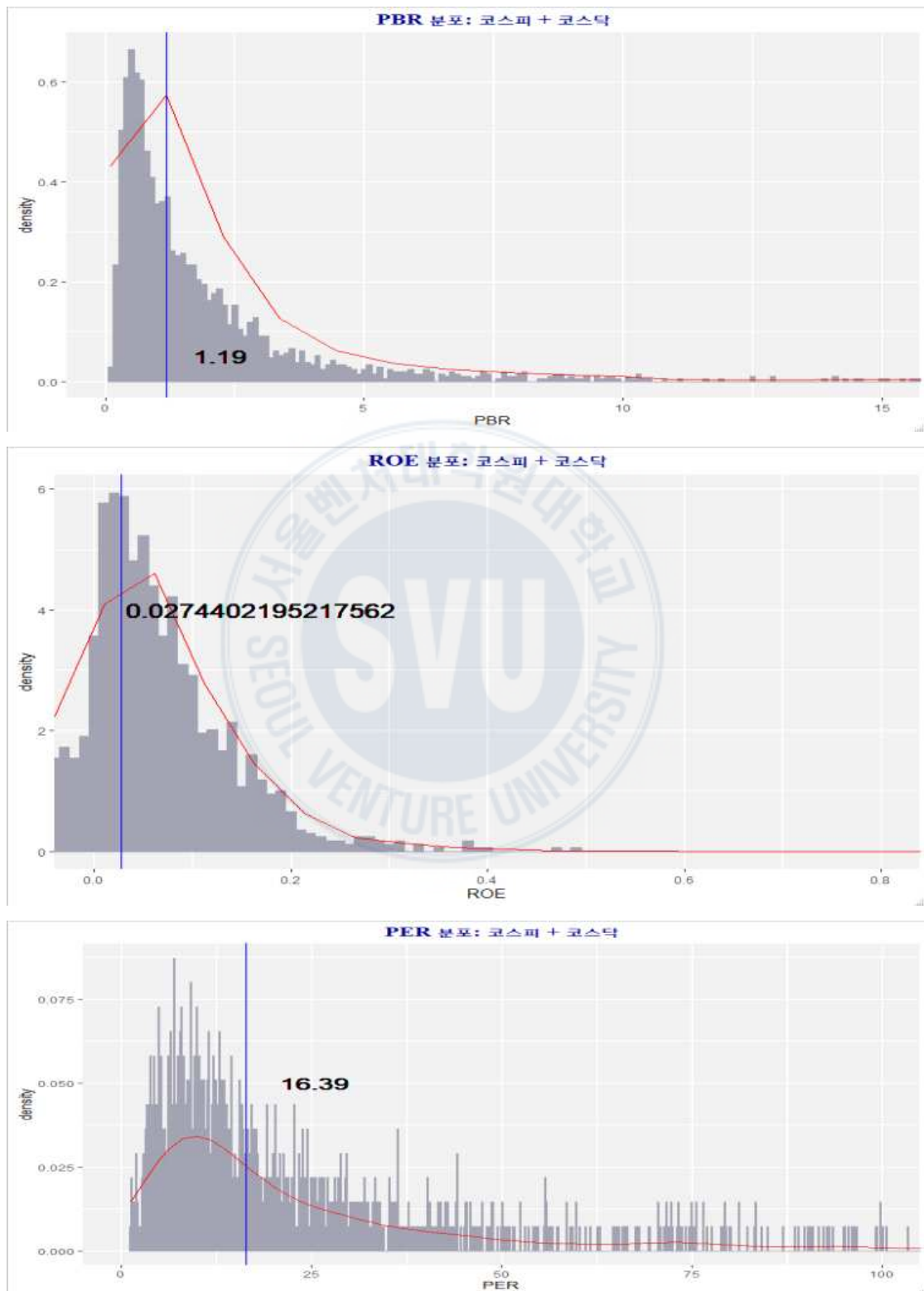


[그림 3-11] 종목선정 재무지표 간의 상관계수

본 연구에서는 수집한 데이터로부터 각 지표의 Rank를 구하고 앞서 투자지표와 같이 상관계수를 살펴보았다. [그림 3-11] 에서 보듯 위에서 투자지표들과 같이 상관계수로 볼 때 같은 양의 관계라고 하더라도 강도의 차이가 있어 동시에 고려할 필요가 있으며 다른 지표들과도 함께 살펴봐야 할 것이다.

지금까지 살펴본 기준 중 투자가치 지표와 재무지표 중 PER, PBR, ROE에 대한 주식시장의 전체 종목 분포를 히스토그램으로 나타내 보면 [그림 3-12]와 같다.

이러한 여러 지표를 반영함으로써 선택하는 종목을 다양화에 반영하여 포트폴리오 구성 시 위험을 분산시킬 수 있다. 이러한 재무지표 나 투자지표들 또는 그들 간의 관계를 투자자는 경제지표, 사회지표 등의 기업 외부 영향 변수들과 연계하여 연구해야 할 부분이다. 본 연구에서는 대외 영향요인과 기업성과나 기업 재무상태에 관한 상관관계 연구는 연구의 범위를 벗어나 다루지 않지만, 향후 연구의 실용성을 높이기 위해서는 다변량 영향 관계 연구로 확장해야 할 필요가 있다.



[그림 3-12] PER, PBR, ROE의 주식시장에서의 분포

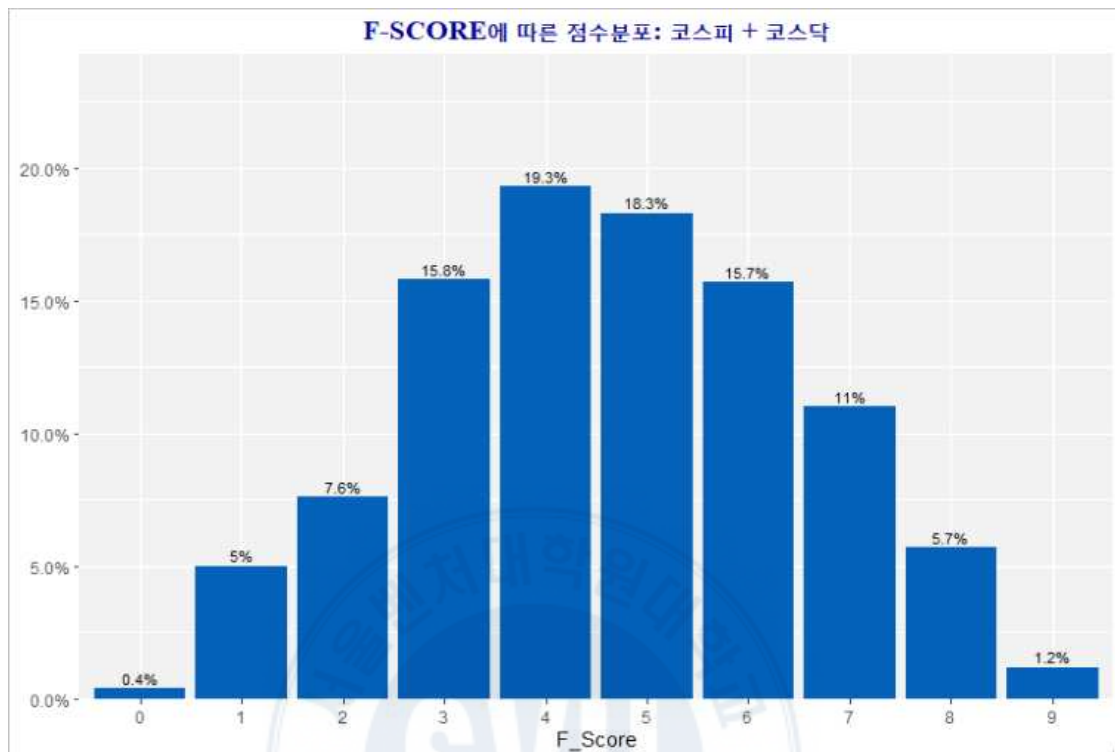
그런데 여러 지표를 동시에 반영하여 종목을 선택 포트폴리오를 구성할 경우, 단순히 이들 지표의 Rank를 점수로 합을 구해 전체 합계 Rank 점수가 높은 순으로 선택해 가는 방법은 위험성이 있다.

포트폴리오에 각 지표를 비중을 정해 적용할 때는 단순 Rank 합계로는 각 지표의 단위나 크기가 달라 이를 포트폴리오 구성 비중으로 가중평균하면 왜곡되는 문제가 있기 때문이다. 이런 경우 정규분포 형태로 정규화하여 반영하거나, 가치지표들을 이용한 적용 규칙을 만드는 것이 필요하다.

학계나 투자전문가들은 재무지표와 투자가치 지표를 조합하여 기업의 우량성에 대하여 투자자들에게 지표로 제공하는데 대표적인 내용을 요약해 보면 Profitability(수익성), Earnings stability (수익의 안정성), Capital structure(기업구조), Growth(수익 성장성), Accounting quality(회계 우량성), Payout/dilution(배당), Investment(투자) (Hsu, Kalesnik, and Kose 2019) 등이다.

예로 F-Score지표(Piotroski and others 2000)는 낮은 PBR 종목 중 재무적으로 우량한 기업을 선정해 투자하는데 이때, 수익성(Profitability), 재무성과(Financial Performance), 운영 효율성(Operating Efficiency)으로 구분해 총 9개의 지표를 선정하고 우수 지표에 1점, 그렇지 않은 지표에 0점을 채점하고 합산하여, 총 0점부터 9점까지의 점수로 종목을 선택한다.

다음의 [그림 3-13]에서는 코스닥 시장과 코스피 시장 기업의 재무제표를 이용하여 2020년도 이전 3개년도를 내용으로 F-SCORE를 구하고 그 주식시장 종목의 점수분포를 히스토그램으로 나타냈다.



[그림 3-13] F-SOCRE에 따른 주식시장 분포

나. 멀티기준 종목선정

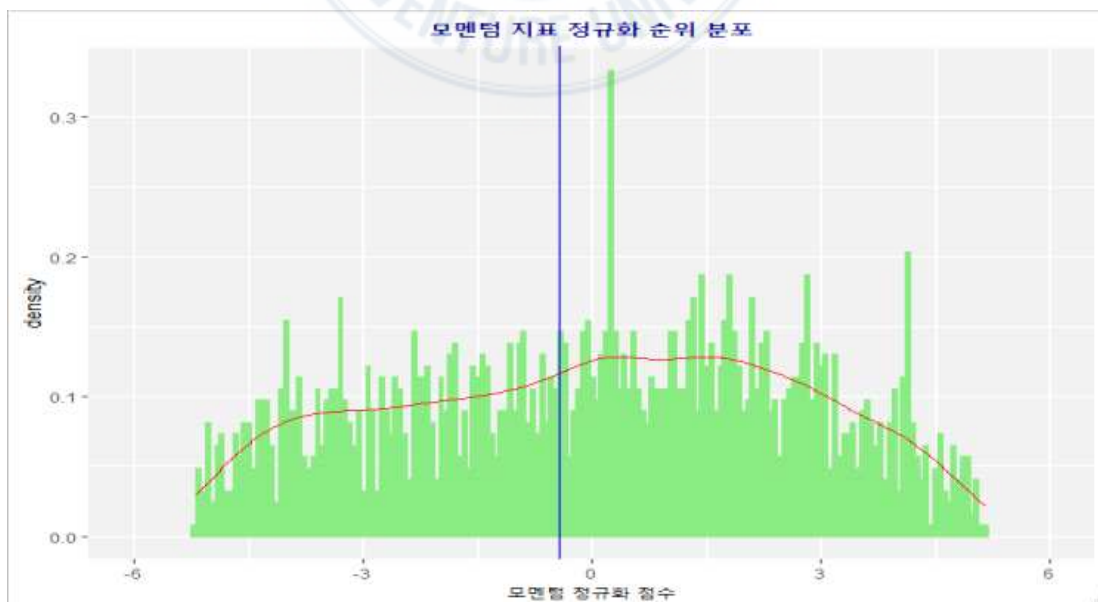
지금까지 주가와 정보를 수집하고 그 데이터로부터 본연구의 모델에 적용하여 변동성이나 포트폴리오 구성비를 구하기 위한 종목선정 기준을 살펴보았다. 그 과정에서 KOSPI 및 KOSDAQ의 특성을 조금은 살펴볼 수 있었다. 앞에서 살펴본 선정 기준들을 사용하여 다음과 같이 종목을 선정하였다. 첫째, 시장의 산업분류별 기초적인 정보를 살펴보기 위한 가격변동을 반영하는 베타(β)를 기초적으로 살펴보았다. 둘째, 수익률의 추세를 고려하기 위한 모멘텀(단, 위험조정은 하지 않는다). 셋째, 기업의 발전성과 시장가치를 반영하기 위하여 금융투자에서 사용되는 투자 가치지표 넷째, 기업의 안정성과 성과를 반영하기 위한 재무 가치지표를 사용하여 종

목을 선정하였다.

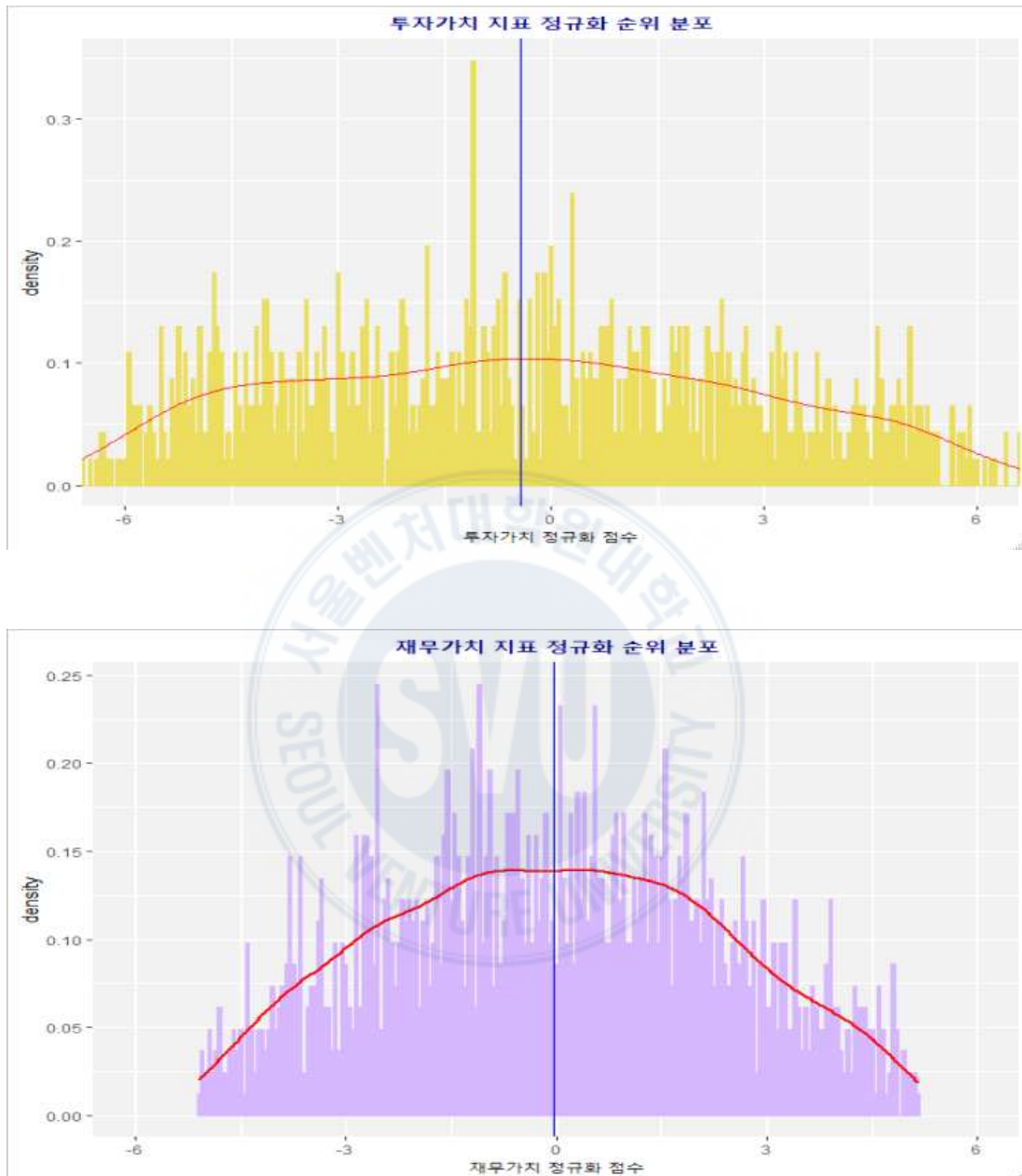
기타 산업군별 가중치나 경제 상황 이슈에 따른 테마종목 구분이나, 새로운 전략 분석 기준 등은 본 연구의 방향이 시계열 데이터 분석과 강화학습의 적용 및 가상데이터 적용에 있기에 사용하지 않고 선정하였다. 실제 투자에서는 기업 외적 변수, 주식시장 외적 변수, 경제-사회-국제 환경변수들의 외부 변화까지 모두 고려하여야 할 것이다.

선정에 사용된 데이터는 2020년 7월 21일 기준으로 이전 2년의 데이터를 기본으로 수집된 데이터를 사용하여, 일정 기간의 변동성 및 시장의 상황만을 고려하고자 하였다.

[그림 3-14] 각 지표를 전체 주식 종목에 적용하여 모멘텀, 투자가치 지표, 재무가치 지표에 정규화를 적용하고 종목선정을 위한 순위(Rank)를 계산한 후 분포를 확인한 것이다. 어느 정도 정규분포 형태를 나타내고 있음을 확인할 수 있다.

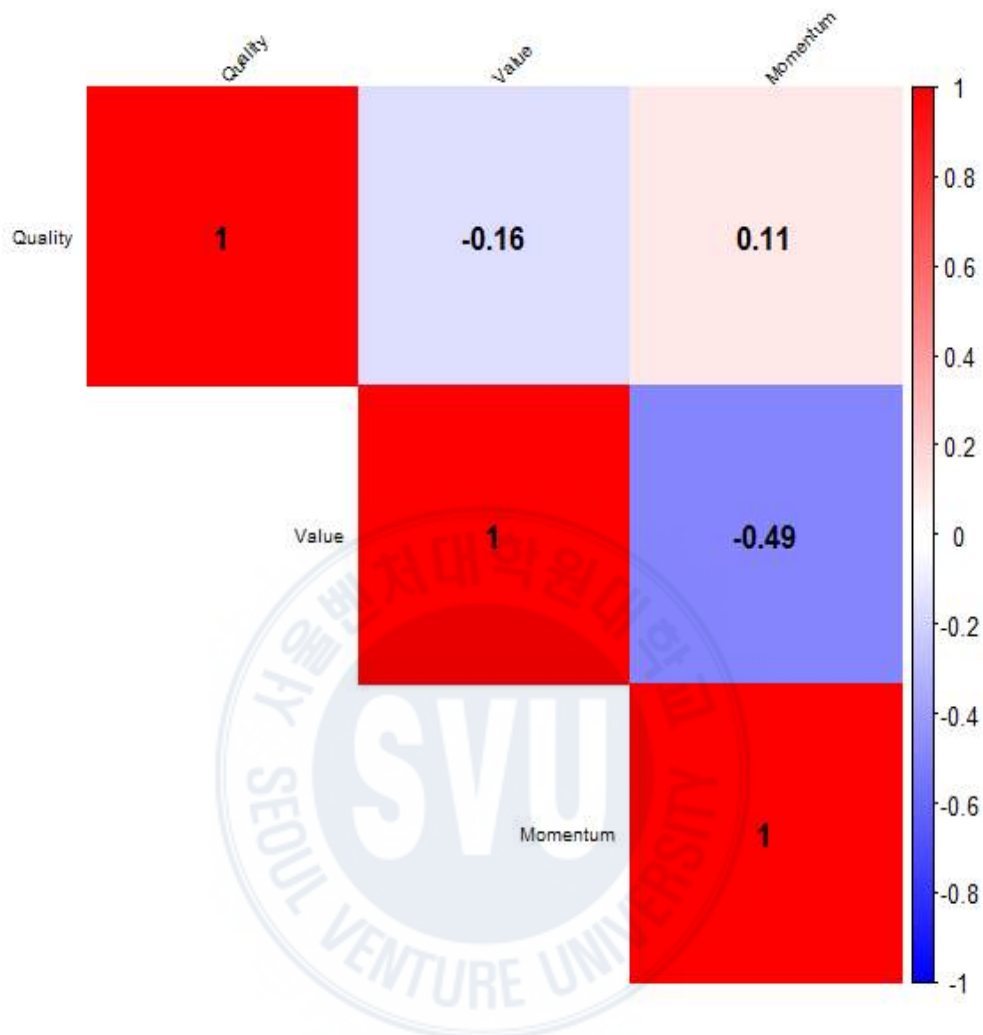


[그림 3-14] 종목선정 지표의 시장종목 순위 분포



[그림 3-14] 종목선정 지표의 시장종목 순위 분포

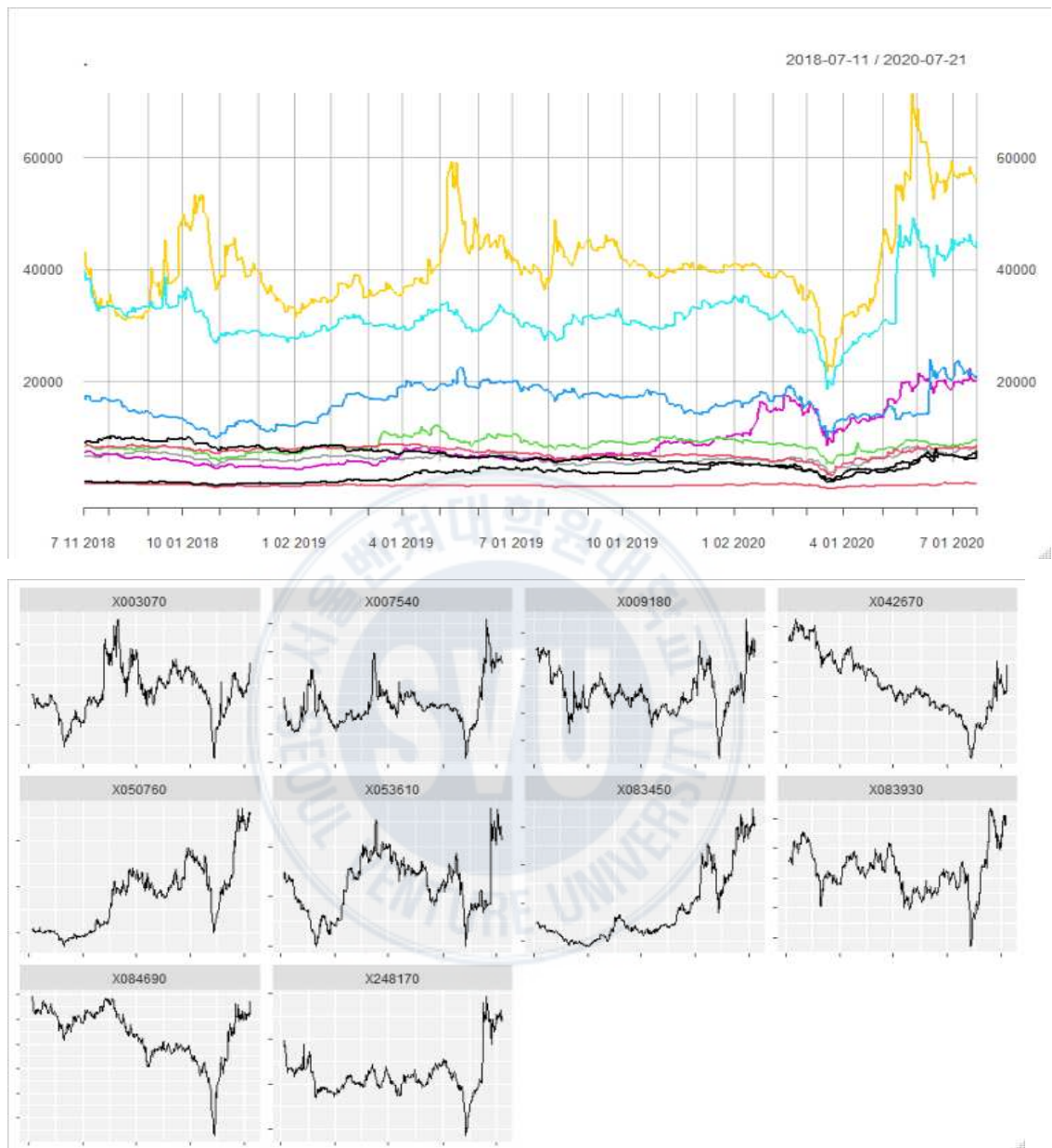
이러한 종목선정 지표에 대한 정규분포 형태는 환경변수를 사용하여 모델에 무작위 생성 변수 Simulation을 반영할 수 있음을 보여 준다.



[그림 3-15] 종목선정 지표 간의 상관계수

이렇게 계산되어진 지표 간의 상관관계를 다음 [그림 3-15]와 같이 확인하였다. 이러한 상관관계나 지표 간의 특성 등이 지표선정에는 중요하지만 본 연구에서는 특별한 비중 전략이나 종목선정 전략을 사용하지 않고 각 지표 간의 비중을 똑같이 설정하여 종목을 선택하였다.

[그림 3-16] 은 종목선정 기준 지표에 따라 상위 10개의 종목을 선정하고 각 선정된 주가 종목의 지난 2년간의 주가를 그래프로 나타내었다.



종목코드	X003070	X007540	X009180	X042670	X050760	X053610	X083450	X083930	X084690	X248170
종목명	코오롱글로벌	샘표	한솔로지스틱스	두산인프라코어	에스폴리텍	프로텍	GST	아바코	대상홀딩스	샘표식품

[그림 3-16] 선정 10종목의 주가변동 그래프

3. 변동성 분석과 포트폴리오 구성

가. 포트폴리오 구성비 기준

포트폴리오의 변동성은 공분산(Covariance)을 사용하며 다음의 (수식 23)과 같이 각 자산의 비중을 계수로 사용하여 각 비중만큼의 공분산에 가중한 값을 사용하여 구할 수 있다.

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} \quad (\text{수식 23}) \qquad w' \Omega w \quad (\text{수식 24})$$

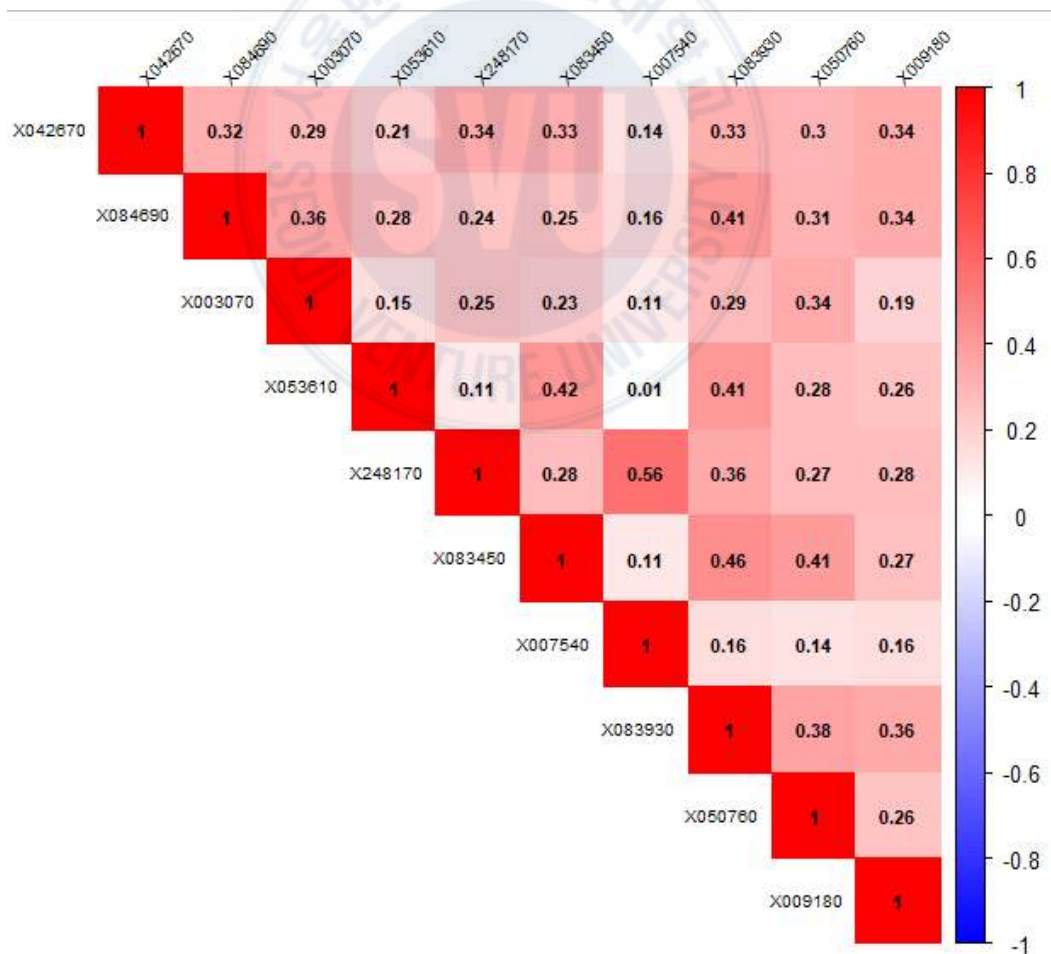
그러나 R이나 Python과 같은 컴퓨터 프로그램을 이용하는 경우는 비중들을 행렬의 (w 와 역행렬 w') 형태로 계산하고 분산-공분산도 행렬의 (Ω)형태로 (수식 24)와 같이 사용하는 것이 일반적이고 편리하기에 포트폴리오 최적화에는 분산-공분산 행렬을 대부분 사용한다.

포트폴리오는 각 종목의 상관관계에서 상관관계가 낮을수록 총 포트폴리오의 변동성이 낮아지는 분산효과를 목적으로 한다. 즉, 위험을 낮춘 상태에서 선정한 종목들을 구성하여 높은 수익률이나 목표한 수익률을 얻고자 하는 것이다.

$$MRC_i = \frac{\delta \sigma_p}{\delta w_i} \quad (\text{수식 25}) \qquad \frac{\Omega w}{\sqrt{w' \Omega w}} \quad (\text{수식 26})$$

이때 포트폴리오를 구성하는 종목들의 비중과 전체 포트폴리오의 위험

(변동성)은 입력과 함수의 관계로 볼 수 있다. 포트폴리오의 한 구성 종목의 비중의 변화량에 따라 전체 포트폴리오의 변동성이 변화하기 때문이다. 이는 다시 말해서 수학적 편미분의 개념으로 포트폴리오의 각 구성 종목의 비중의 변화량에 대한 한계 위험기여도(MRC: Marginal Risk Contribution)를 구할 수 있다는 뜻이며, 이에 다시 비중을 곱해주어 위험기여도(RC: Risk Contribution)를 계산해 낼 수 있음을 말한다. (수식 25)는 이러한 한계위험도에 대한 식을 나타내며, 이를 다시 행렬식으로 표현하면 (수식 26)과 같다. 아래 [그림 3-17]은 본 연구에서 선정한 10개 종목 간 상관계수이다.



[그림 3-17] 선정 10개 종목의 상관계수

이러한 한계위험도와 기여위험도를 이용하여 우리는 포트폴리오의 위험을 계산할 수 있고 그에 따라 여러 전략을 세울 수 있게 된다. 앞에서 언급하였듯, 각 구성 종목의 비중과 총 위험의 함수식으로 포트폴리오와 구성 비중과의 관계를 설명할 수 있었다. 이는 원하는 전략에 따라 제약조건을 사용하여 함수식의 최적 해를 구할 수 있다는 뜻이 된다. 그러나 다변량 함수식을 손으로 푸는 것은 어려운 일이며 비효율적이다.

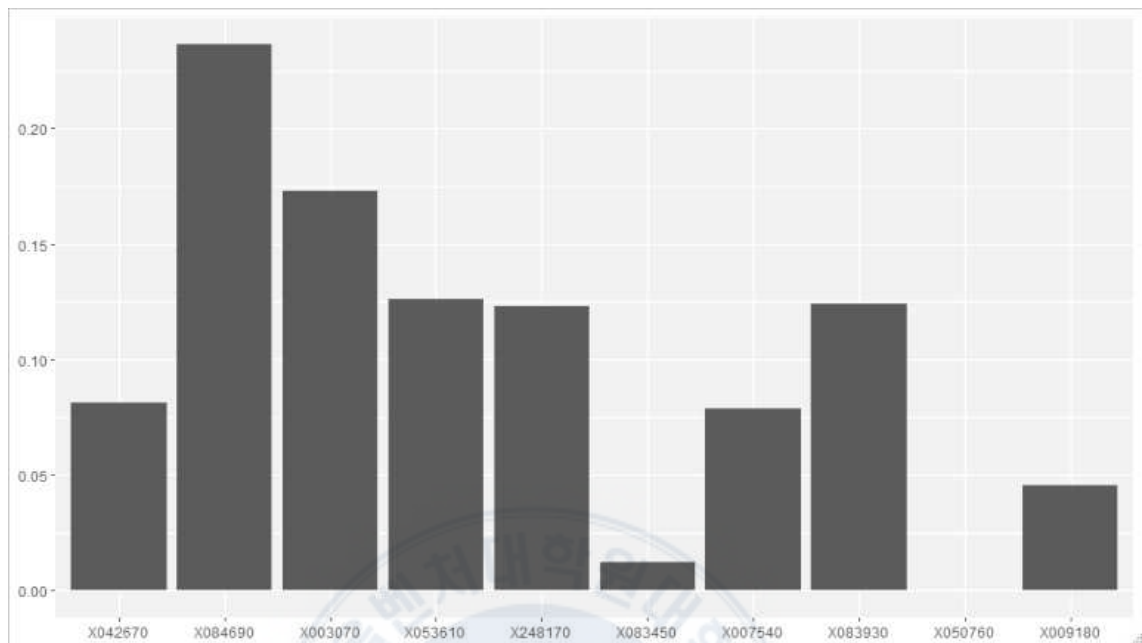
본 연구에서는 R 프로그램의 nloptr 패키지의 slsqp() 함수, quadprog 패키지의 solve.QP() 함수, RiskPortfolios 패키지의 optimalPortfolio() 함수 등을 사용하였다. 이 같이 Python 등의 오픈소스 프로그래밍 언어에는 여러 공개 모듈들이 있으며 이를 활용하면 쉽게 계산할 수 있다.

a. 최소변동성 기준(Minimum Variance Portfolio)

총 투자 비중을 1로 놓았을 때, 포트폴리오 변동성이 최소가 되게 하면서 동시에 각 투자 종목의 비중이 0 이상 그리고 모든 비중의 합이 1 이하가 되는 제약조건을 설정한다. 선정된 10개 종목을 R 프로그램의 slsqp() 함수 사용하여 최적 해를 구하여 보면 최소변동성은 0.0003308396 이 나오며 각 비중은 다음 <표 3-1>와 [그림 3-18] 과 같다.

<표 3-1> 최소변동성 기준 10종목의 비중

종목	X0 42670	X0 84690	X0 03070	X0 53610	X2 48170	X0 83450	X0 07540	X0 83930	X0 50760	X0 09180
종목명	코오롱글로벌	샘표	한솔로지스틱스	두산인프라코어	에스폴리텍	프로텍	GST	아바코	대상홀딩스	샘표식품
비중	0.0812	0.2364	0.1730	0.1260	0.1232	0.0123	0.0785	0.1240	0.0000	0.0454

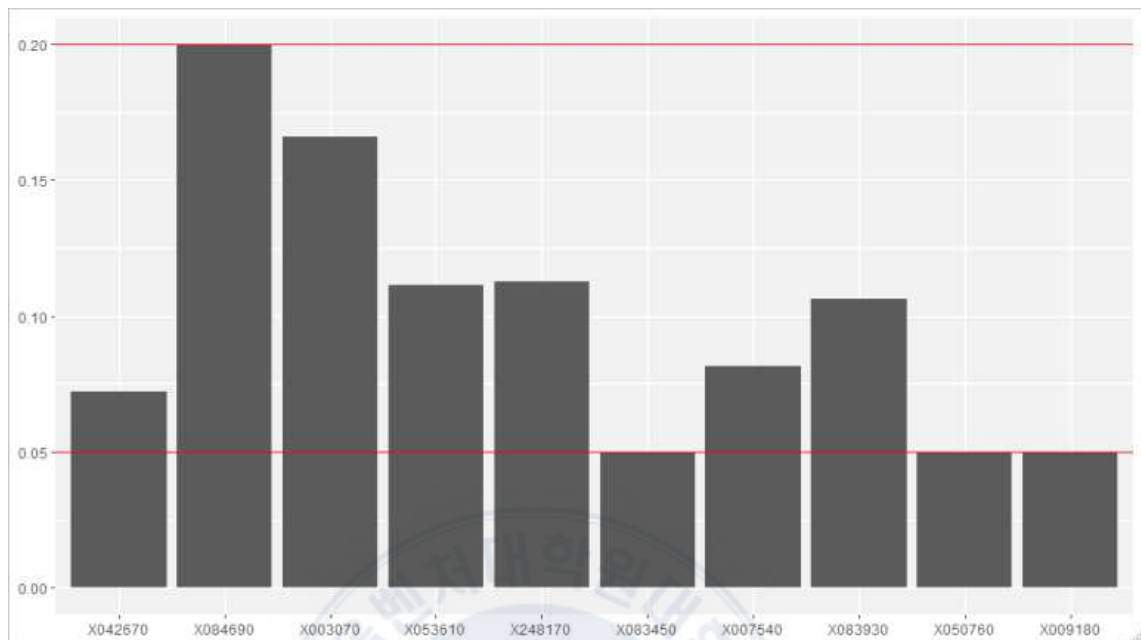


[그림 3-18] 최소변동성 기준 10종목의 비중

그런데 비중이 0인 종목이 발생하였다. 이런 경우 투자전략 상 그 종목을 뺄 수 없는 문제가 있을 수 있다. 이에 강제로 제약조건을 더하여 다음과 같이 최소 5% ~ 최대 20% 비중 제약조건을 추가하여 최적 해를 구할 수 있다. 각 비중은 다음 <표 3-2>와 [그림 3-18] 과 같다.

<표 3-2> 최소변동성 기준 10종목의 비중(비중 제약조건)

종목	X042670	X084690	X003070	X053610	X248170	X083450	X007540	X083930	X050760	X009180
종목명	코오롱글로벌	샘표	한솔로지스틱스	두산인프라코어	에스폴리텍	프로텍	GST	아바코	대상홀딩스	샘표식품
비중	0.0722	0.2000	0.1660	0.1115	0.1126	0.0500	0.0815	0.1062	0.0500	0.0500



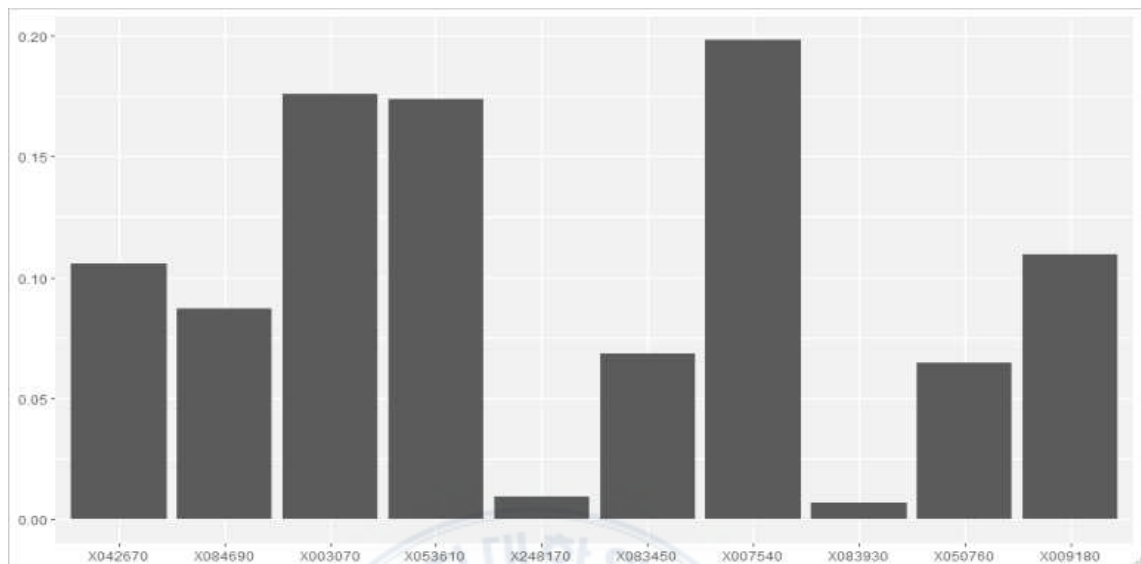
[그림 3-19] 최소변동성 기준 10종목의 비중(비중 제약조건)

b. 최대변동성 기준

최소변동성 기준에서와 반대로 포트폴리오의 변동성 총합이 최대가 되도록 각 종목의 비중을 계산하는 것을 말하며 이때의 비중은 다음 <표 3-3>과 [그림 3-19]와 같다.

<표 3-3> 최대변동성 기준 10종목의 비중

종목명	종목코드	비중
코오롱글로벌	X042670	0.1057
샘표	X084690	0.0873
한솔로지스틱스	X003070	0.1757
두산인프라코어	X053610	0.1740
에스폴리텍	X248170	0.0092
프로텍	X083450	0.0687
GST	X007540	0.1982
아바코	X083930	0.0067
대상홀딩스	X050760	0.0649
샘표식품	X009180	0.1096



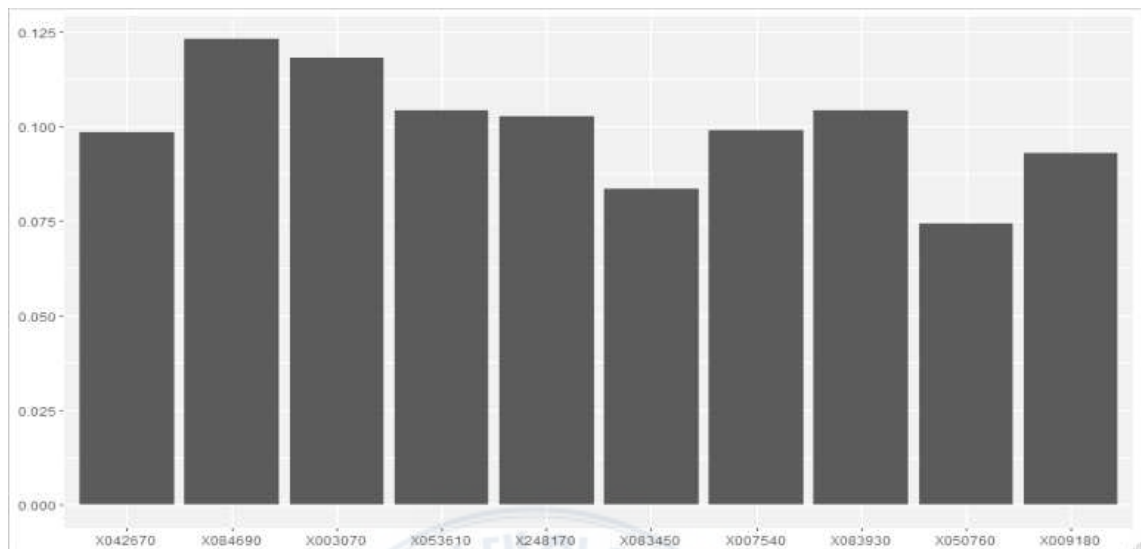
[그림 3-20] 최대변동성 기준 10종목의 비중

c. 위험균형 기여위험률 평준화 기준

다른 비중 제약조건으로 위험균형 기여위험률 평준화 기준을 주는 방법이 있는데, 다른 말로는 리스크패리티(RiskParity)(Qian 2011)라고 한다. 앞에서 살펴본 한계위험기여도(MRC)를 이용하여 각 구성 종목의 위험기여도를 평준화하는 것을 말한다. 특정 종목의 가격변동에 전체 포트폴리오가 과도하게 영향받는 것을 피하기 위한 전략이다.

<표 3-4> 위험균형 기여위험률 평준화 10종목의 비중

종목	X042670	X084690	X003070	X053610	X248170	X083450	X007540	X083930	X050760	X009180
종목명	코오롱글로벌	샘표	한솔로지스틱스	두산인프라코어	에스폴리텍	프로텍	GST	아바코	대상홀딩스	샘표식품
비중	0.0985	0.1230	0.1180	0.1041	0.1025	0.0835	0.0990	0.1042	0.0744	0.0929



[그림 3-21] 위험균형 기여위험률 평준화 10종목의 비중

이처럼 분석가나 투자자의 투자전략에 따라서 여러 제약조건과 계산 방법, 주기 등을 다양화할 수 있다. 예를 들어 특정 기간별 변동성과 공분산을 재계산하고 비중을 갱신할 수 있다. 또는 특정 목적을 위하여 산업군이나 종목을 지정하여 일정 비중을 유지할 수도 있다. 경우에 따라서는 선정된 종목 일부분만을 전략적 기준에 따라 비중을 재계산하거나 제약을 가할 수도 있다.

이러한 비중 계산을 통한 전략뿐만 아니라 새로운 접근 방식도 있을 수 있다. 예로 위험예산 포트폴리오, 전체 주가지수 추세를 이용한 인덱스 포트폴리오 등 여러 전략이 있으며, 실전 투자에서는 많이 활용되고 있다.

나. GARCH를 통한 변동성 예측

지금까지 살펴본 고전적 투자방법론은 모두 역사적 데이터를 바탕으로 과거로부터 현재를 평가하고 투자의 방향을 결정하는 방식이었다.

그러나 컴퓨터 연산능력의 발전과 새로운 예측을 위한 모델화 방법들이 발전하면서부터 역사적 데이터로부터 정형화된 패턴과 모델을 개발, 앞으로의 가능성을 예측하여 투자하는 방법이 발달하였다.

본 연구에서는 먼저 앞에서 설명한 GARCH 모델을 이용한 방법론을 적용해 보고 다음으로 RNN(Recurrent Neural Network)를 이용한 시계열 예측을 적용해 보았다.

GARCH 모델을 다루기 위해서는 먼저 과거의 데이터로부터 정상성(stationality) 시계열 모형의 적합성 테스트가 필요하다. 이를 위하여 정상성은 dickey-fuller test를 시계열 모형의 적합성을 판정하기 위하여는 LM(Ljung-Box test)을 실시하였다. Box와 Pierce(1970년)가 ARIMA 모형의 잔차들로 부터 검정 통계량 Q를 (수식 27)과 같이 개발하였다.

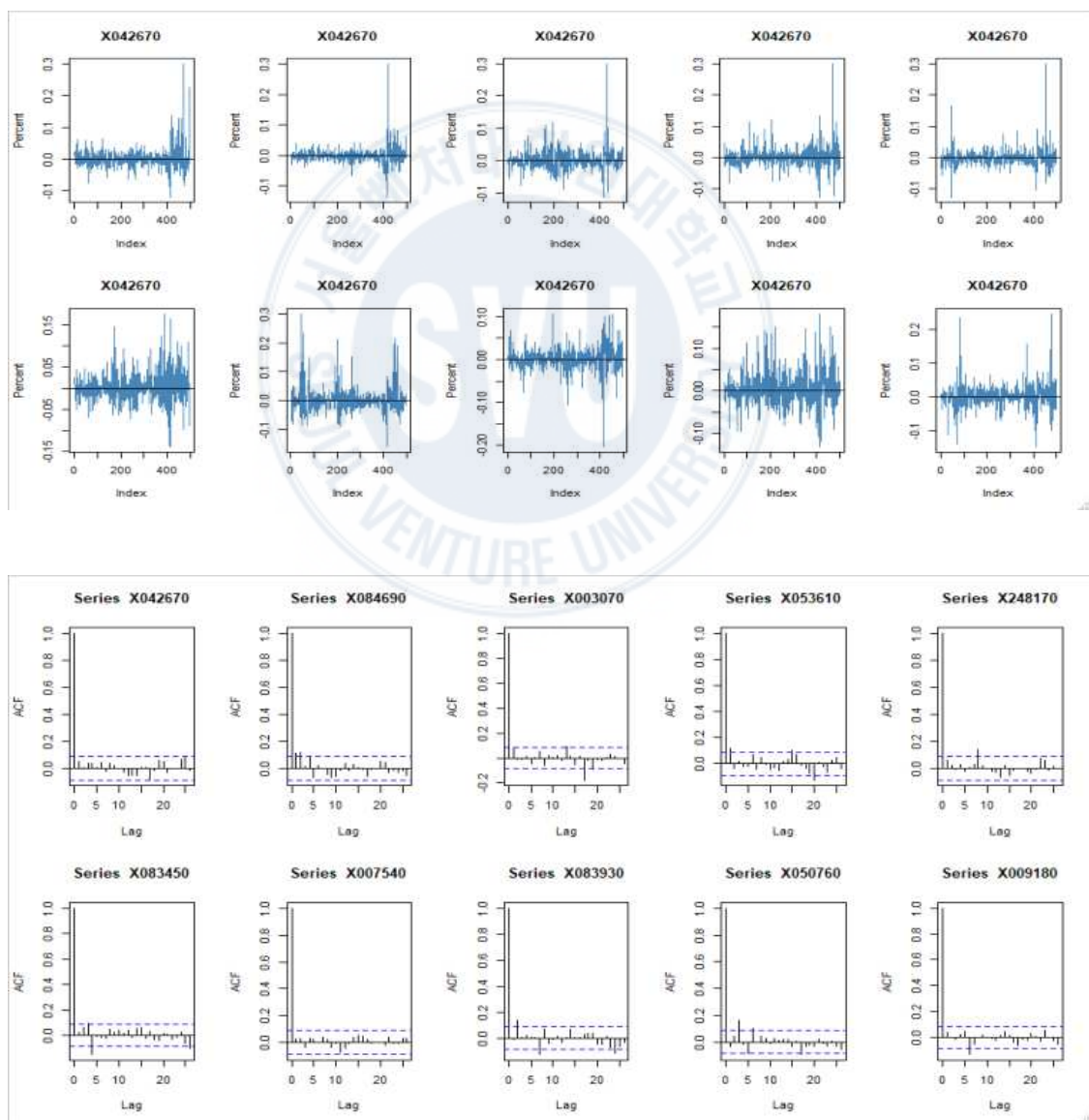
$$Q = n(\hat{r}_1^2 + \dots + \hat{r}_k^2) \quad (\text{수식 27})$$

이후 Ljung과 Box(1978년)가 이를 개선하여 (수식 28)을 고안하였다.

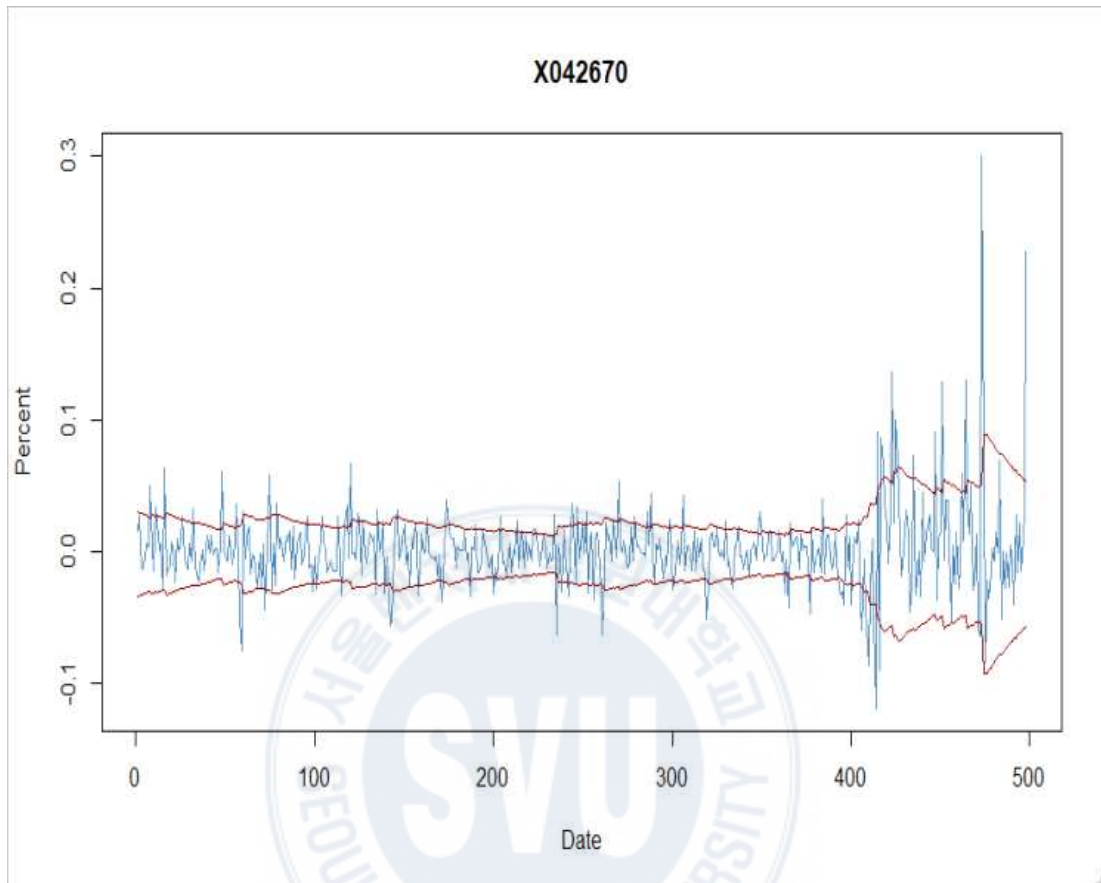
$$Q_* = n(n+2)\left(\frac{\hat{r}_1^2}{n-1} + \dots + \frac{\hat{r}_k^2}{n-k}\right) \sim \chi^2(k-p-q) \quad (\text{수식 28})$$

본 연구의 예시 GARCH(1, 1) 모형은 수익률을 사용하기 위하여 log 수익률을 이용하여 1차분 한 데이터를 사용하였다. 수익률에 대하여 ACF(자기상관함수)를 분석해보면 10종목 모두 자기상관함수의 코릴로그램(Correlogram)이 이와 유사함을 볼 수 있다. 그러나 자기상관함수(ACF)와 PACF(편자기상관함수)의 코릴로그램만 가지고 판단하기에는 부족함

이 있어 확장자기상관함수(EACF:Extended Autocorrelation Function)를 사용하여 모형의 차수를 계산하였다. 이는 R 프로그램 package TSA의 EACF 함수를 사용하여 구할 수 있다. [그림 3-21] 10종목에 대한 log 수익률과 선택한 10종목에 대한 log 수익률을 이용하여 1차분 한 변동성 ACF 코릴로그램이다.



[그림 3-22] 10종목에 대한 log 수익률과 ACF 그래프



[그림 3-23] X042670(두산인프라코어) 종목에 대한 log수익률과 GARCH 모델

이때 시계열모델 적합성에 대한 분석의 예로 10개 선택종목 중 X042680에 대하여 살펴보면, [그림 3-22] 10종목에 대한 log 수익률과 ACF 그래프와 같은 모형이 나오는데 흐린 진동하는 그래프는 실제 데이터를 나타내고 굵은 실선의 범위는 모형을 의미한다. 이와 같은 GARCH(1, 1) 모형을 식으로 표현하면 다음 (수식 29)와 같다.

$$\sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2 + \epsilon_t \quad (\text{수식 29})$$

즉, 한 시점 전의 수익률과 변동성을 통해서 현재의 변동성을 구하는 모형이다. 이에 대한 GARCH 모형을 R 프로그램의 package fGarch 함수를 사용하여 최적 계수의 해를 구해보면 계수(Coefficient)는 다음과 <표 3-5>와 같다.

<표 3-5> X042670(두산인프라코어) 종목 GARCH 모형의 계수(Coefficient)

Coefficient(s):			
mu	omega	alpha1	beta1
-1.6513e-03	8.9970e-07	6.2731e-02	9.4722e-01

이때, 각 계수의 유의성은 다음 <표 3-6>과 같다. 결과를 보면 절편과 상수항은 유의하지 않으나 수익률에 대한 계수와 변동성에 대한 계수는 유의하다고 확인되었다.

<표 3-6> X042670(두산인프라코어) 종목 GARCH 모형의 유의성

Error Analysis:				
	Estimate	Std. Error	t value	Pr(> t)
mu	-1.651e-03	9.790e-04	-1.687	0.0917
omega	8.997e-07	2.754e-06	0.327	0.7439
alpha1	6.273e-02	1.345e-02	4.664	3.1e-06 ***
beta1	9.472e-01	1.210e-02	78.284	< 2e-16 ***

또한, 앞에서 설명한 이러한 모델의 시계열 적합성을 판단하기 위한 테스트로 LM(Ljung의 Box(1978년) 테스트의 결과를 보면 다음과 같다. 이때 귀무가설은 ‘이분산성이 없다’이며, 대립가설은 ‘이분산성이 있다.’이다. <표3- 7>의 결과에서 볼 수 있듯이, 모든 잔차의 유형에서 귀무가설을 기각할 수 없다. 따라서 시계열 모델로서 유의하다. 또한 Jarque-Bera test의 모형의 잔차의 정규분포성도 확인되었다.

<표 3-7> X042670(두산인프라코어) 종목 GARCH 모형의 Ljung-Box test

Standardised Residuals Tests:				
			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	297.0998	0
Shapiro-Wilk Test	R	W	0.9540972	2.523068e-11
Ljung-Box Test	R	Q(10)	4.668855	0.9121736
Ljung-Box Test	R	Q(15)	5.532828	0.9865731
Ljung-Box Test	R	Q(20)	14.01494	0.8297375
Ljung-Box Test	R^2	Q(10)	7.126678	0.7134341
Ljung-Box Test	R^2	Q(15)	8.265225	0.9127097
Ljung-Box Test	R^2	Q(20)	9.171049	0.9808401
LM Arch Test	R	TR^2	8.048087	0.7813613

다만 잔차의 정규분포성을 확인하는 Shapiro-Wilk Test의 값이 귀무가설을 기각하였으나, 주가의 변동성 예측모델의 잔차의 정규성을 확보한다는 것은 극히 어려운 일이다. Jarque-Bera test과 Shapiro-Wilk Test 판정이 일치하지 않는 경우는 많다.

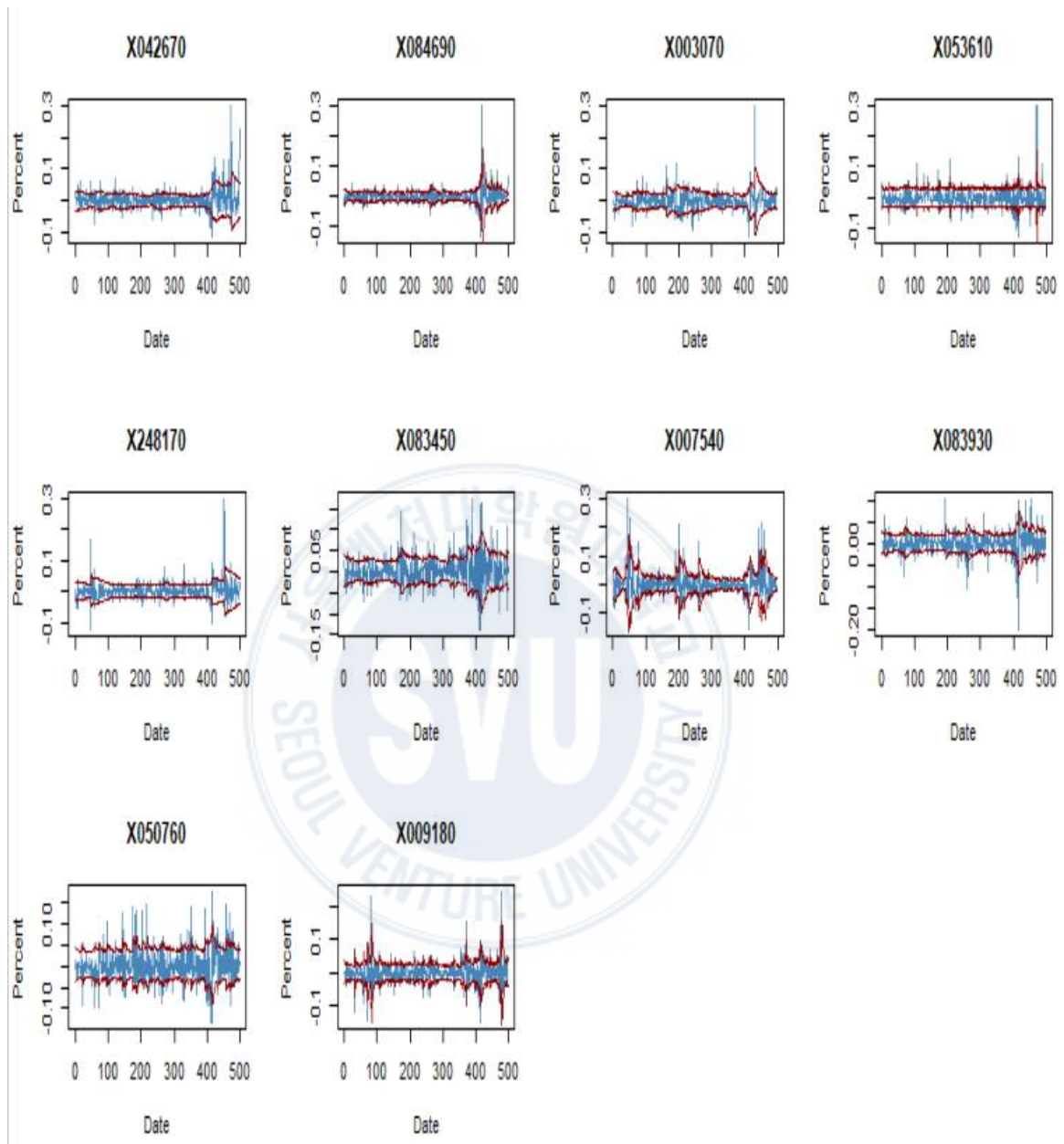
다음의 (수식 30)은 이러한 모형으로부터 앞으로 n 일 후의 평균 변동성을 예측하는 공식이다. V 는 장기평균분산을 뜻한다. 이를 n 차 일의 변동성의 기댓값으로 바꿔 그 평균을 구하여 유도한다.

$$\sigma_{average}^2(t, t+n) = V + (\sigma_t^2 - V) \frac{(\alpha + \beta)1 - (\alpha + \beta)^n}{n(1 - \alpha - \beta)} \quad (\text{수식 30})$$

$$: V = \frac{\omega}{1 - \alpha - \beta}$$

$$: \sigma_{t+1}^2 = V + \alpha(X_{t-1}^2 - V) + \beta(\sigma_{t-1}^2 - V)$$

이와 같은 과정을 통해 10개의 종목의 수익률의 변동성을 예측하는 모델을 생성한다. 다음의 [그림 3-23]는 나머지 9개의 변동성 GARCH 모델이다.



종목명	X003070	X007540	X009180	X042670	X050760	X053610	X083450	X083930	X084690	X248170
종목명	코오롱글로벌	샘표	한솔로지스틱스	두산인프라코어	에스폴리텍	프로텍	GST	아바코	대상홀딩스	샘표식품

[그림 3-24] 10개 종목에 대한 log 수익률과 GARCH 모델

이렇게 구해진 GARCH 모델을 통해서 예측 변동성을 각각의 선택 종목에 대해서 구할 수 있다. 이를 앞에서의 상관계수를 이용하여 공분산을 계산해 낼 수 있다. 즉, 과거의 상관관계와 앞으로의 분산의 변화를 통해서 예측 공분산의 관계를 구할 수 있고 이를 포트폴리오의 최적 비중 예측에 사용한다. 상관계수와 공분산의 관계는 다음과 같다.

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \quad (\text{수식 31})$$

$$|Cov(X, Y)| \leq sd(X)sd(Y) \quad (\text{수식 32})$$

$$Corr(X, Y) \equiv \frac{Cov(X, Y)}{sd(X) \cdot sd(Y)} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}; \quad (\text{수식 33})$$

$$Cov(X, Y) = Corr(X, Y) \times sd(X) \cdot sd(Y) \quad (\text{수식 34})$$

현재는 다변량 GARCH 모형이 개발되어 여러 분야에서 연구되고 있다. BEKK 모형 등과 같은 다변량을 입력으로 사용하는 모델에만 아니라 정상성이나 추세에 연관한 모델들 더 나아가서 변동성의 국면 자체의 변화를 고려하여 상승충격과 하락충격 등의 특정한 외부요인으로 인한 비대칭적 다변량 GARCH 모형 등도 연구되고 있다.

본 연구에서는 이러한 분산예측치를 이용하여 상관계수와 관계이용 공분산을 일정 일자마다 예측하여 최소변동성 포트폴리오의 비중을 계산하여 백 테스트를 시행하였다.

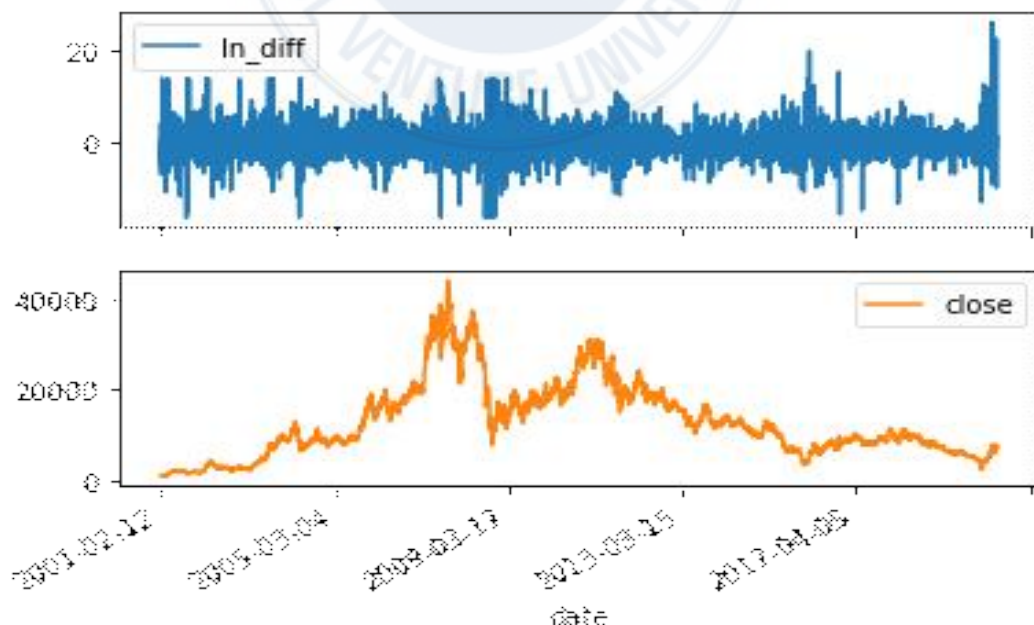
나. RNN을 이용한 예측

자기상관성을 고려한 변화예측이 Deep-Learning의 Neural Network와

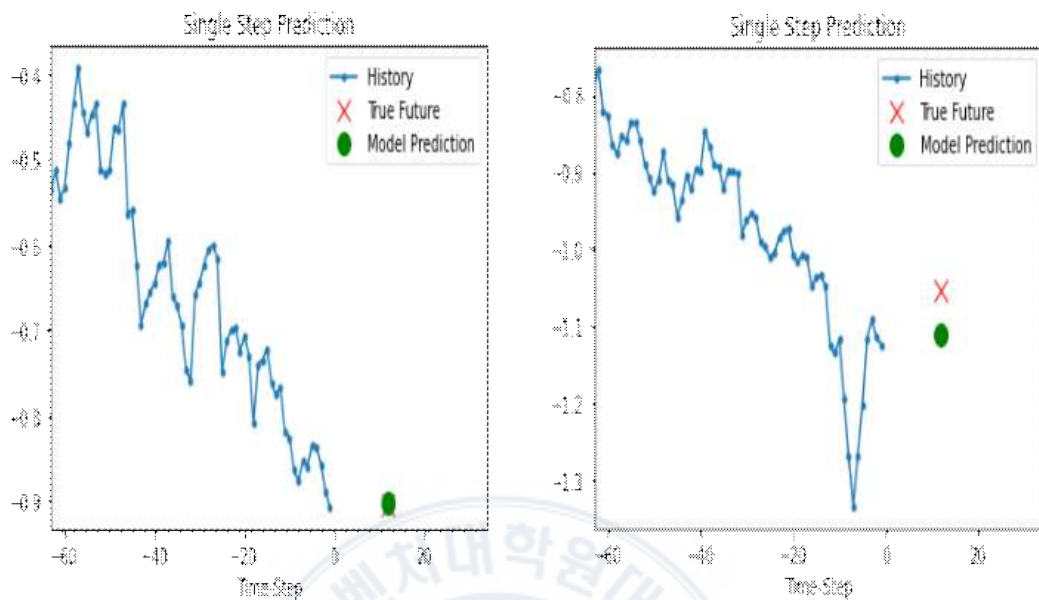
만나 비선형적인 방법으로 가격의 변화를 예측하는 모델이 RNN이다. 본 연구에서는 주식시장의 평균 1달분 영업일인 20일을 예측하고자 하는 기간으로 보고 입력 데이터는 10개 선택 종목의 log 수익률을 이용한 1차분 데이터를 입력으로 사용하였다. 종목의 상한가, 하한가, 거래량 등을 입력으로 함께 사용한 모델까지 고려하였다.

모델의 훈련은 전체 데이터를 train과 validation으로 나누고 1 Step의 훈련에 250일(평균 1년의 주식시장 업무일)로 Sliding Window 하여, 입력으로 사용하여 5일(1주)을 1 Step으로 4 Step을 예측하여 20일을 예측하는 모형을 기본으로 입력에 변화를 주었다.

다음의 [그림 3-24]은 선택된 종목 X042670의 log 수익률 1차분 데이터와 종가의 변화를 보여주며, [그림 3-25]는 훈련 후 250일의 과거 데이터에 대한 1 step의 20일 후 예측을 보여주고 있다.

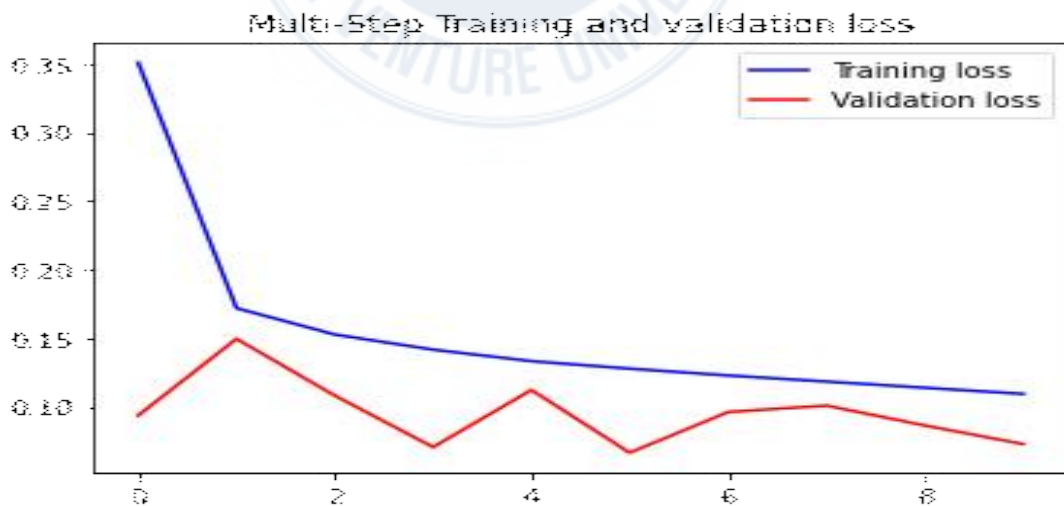


[그림 3-25] X042670(두산인프라코어) 종목에 대한 log 수익률과 종가



[그림 3-26] X042670(두산인프라코어) 종목에 대한 1 Step의 예측 2개

[그림 3-26] X042670 종목에 대한 1Step의 예측 2개는 모델의 학습에 따라 각 모델이 loss 예측 차이가 줄어들고 있음을 알 수 있다.



[그림 3-27] X042670(두산인프라코어) 종목에 대한 LSTM 모델 학습의 Loss 변화 위의 훈련과정을 통한 모델의 Validation Set의 결과 2개를 나타낸 그래프는 아래와 같다.

history: 250, step: 4 (day5) = 4*5 = 20 pred: 20 LSTM : 8, [ln_diff,close]

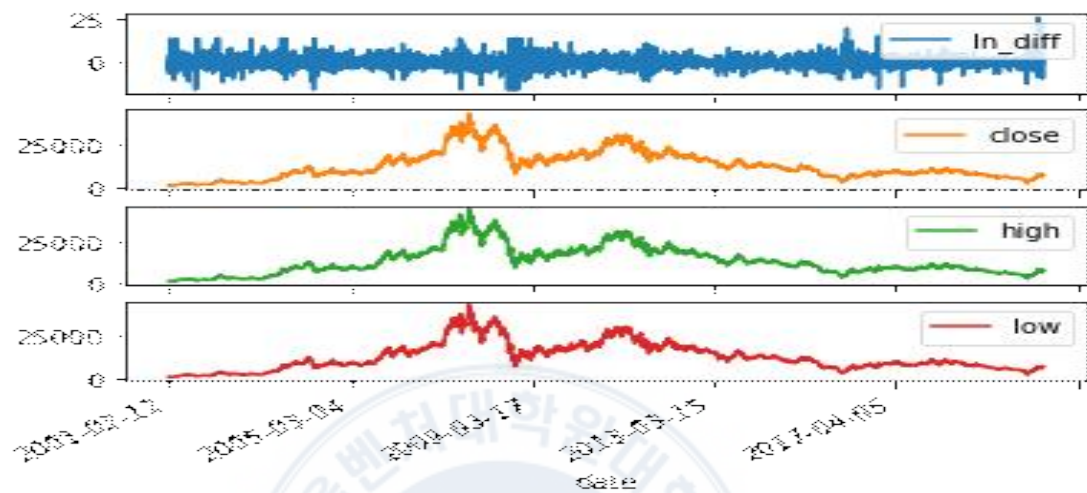


[그림 3-28] X042670(두산인프라코어) 종목에 대한 LSTM 모델의 예측

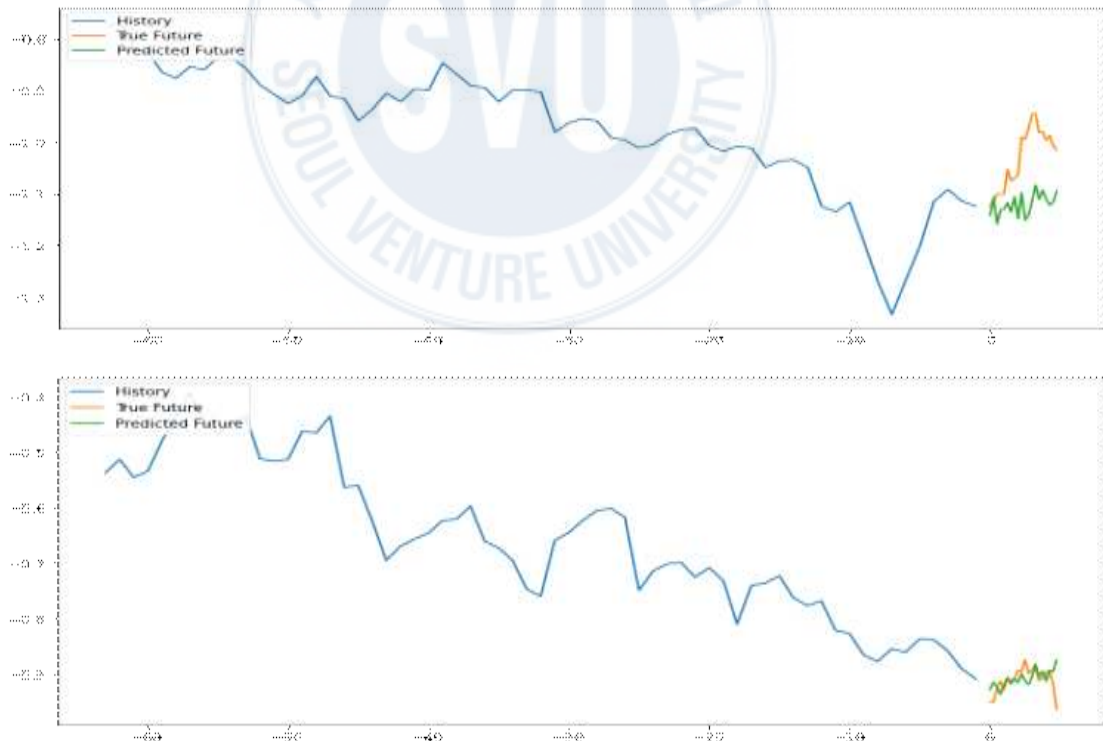
다음 [그림 3-28] X042670 종목에 대한 LSTM 모델의 예측은 입력 및 훈련 세팅에 변화를 준 경우이다. 모든 경우 방향과 추세는 유사하다. 이는 주식 자체의 데이터 외의 요소에 대한 정보를 입력으로 사용하지 않았기 때문으로 보인다. 주가 정보 자체의 변동성으로 앞에 1 Step 1일 예측의 경우와 달리 일정 기간으로 예측이 늘어나면 GARCH 모델과 같이 특정한 추세와 범위의 예측 즉, 기대치나 분산에 대한 정보 확보를 할 수 있다. 반면 예측 정확성과 신뢰도는 떨어지는 것을 알 수 있다.

본 연구의 백 테스트에서는 10일 단위로 하루 앞을 예측한 주가를 이용

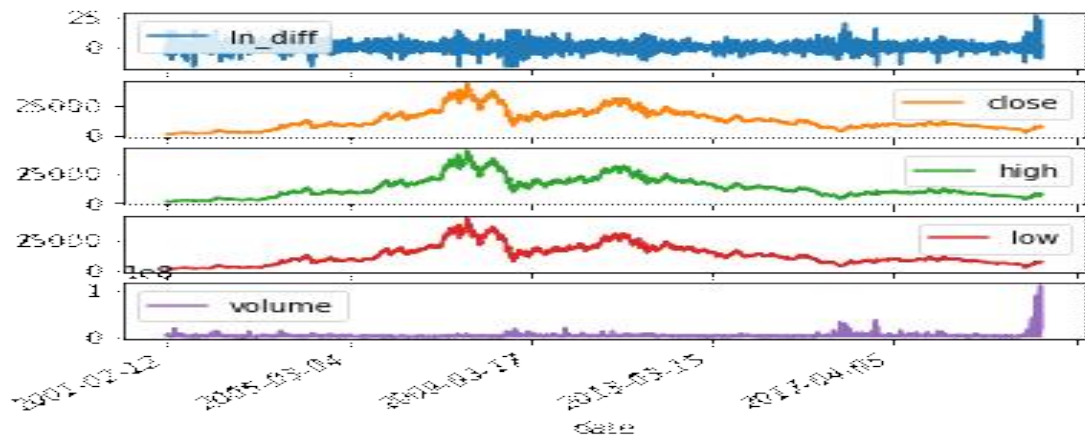
하여 포트폴리오의 분산을 구하고 이를 통해서 포트폴리오 구성비를 구하였다.



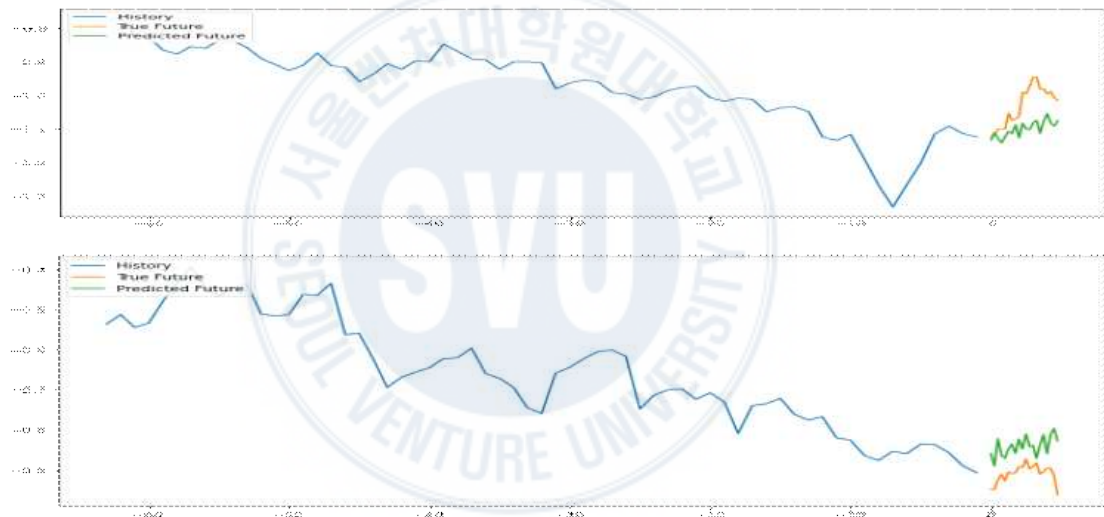
history: 250, step: 4 (day5) pred: 20 LSTM : 8, [ln_diff,close, high, low,volume]



[그림 3-29] X042670(두산인프라코어) 종목에 대한 다변량 LSTM 모델의 예측



history: 250, step: 4 (day5) = 4*5 = 20 pred: 20 LSTM : 8, [ln_diff,close, high, low]



[그림 3-29] X042670(두산인프라코어) 종목에 대한 다변량 LSTM 모델의 예측

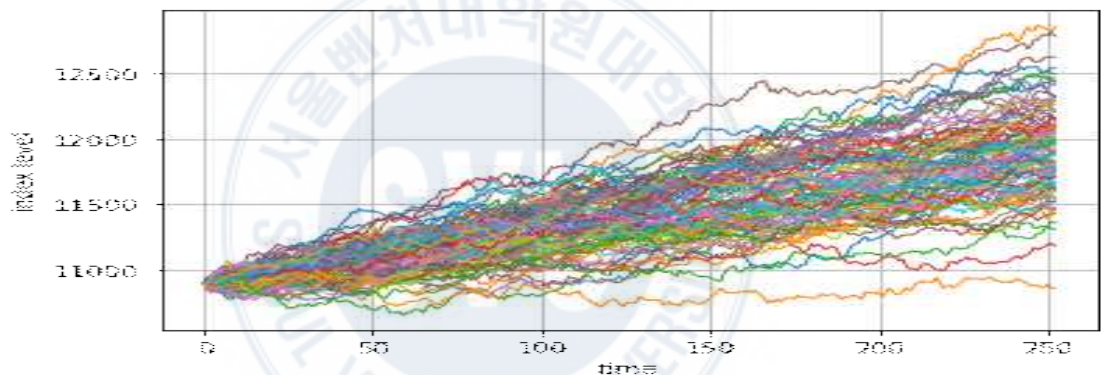
4. GBM을 이용한 DQN

가. GBM Monte Carlo Simulation

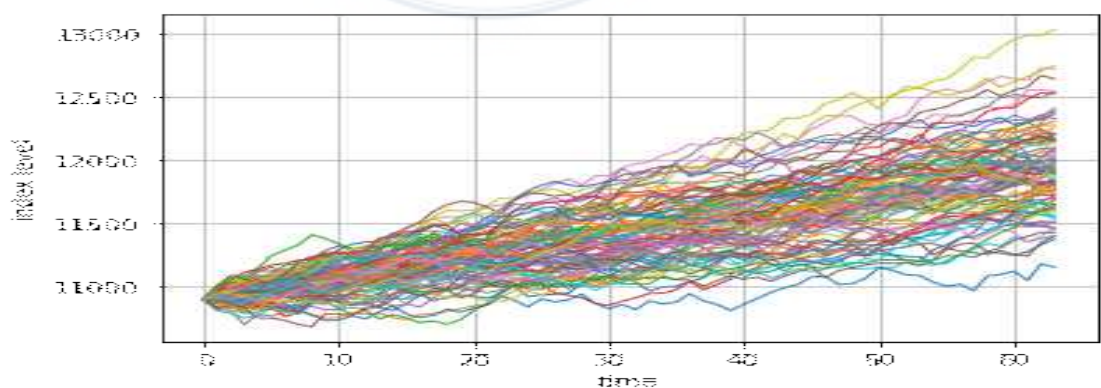
이론적 배경에서 설명한 GBM Monte Carlo Simulation을 통해 DQN의 입력으로 사용하기 위하여 2가지 경우의 데이터를 생성하였다.

Simulation 데이터의 생성은 전체 데이터 중 2년간의(527일) 데이터를 사용 3개월(63일)을 step으로 하는 Window Sliding으로 나누고 무작위 선택하고 한다. 각 step에서 2달(63일)의 데이터 생성 또는 1년(252일)의 데이터를 GBM Monte Carlo Simulation으로 생성하였다.

다음 [그림 3-29]은 X042670(두산 인프라코어)의 한 window에서 100번의 무작위 생성 시 정규분포 형태의 랜덤한 데이터가 형성되는 것을 보여주고 있다. 즉, 확률적으로 적정한 정규분포를 이루는 데이터가 생성됨을 보여주고 있다.



(252일의 Simulation 데이터를 생성)



(63일의 Simulation 데이터를 생성)

[그림 3-30] X042670(두산인프라코어) 종목의 1step의 GBM MC simulation 100회 생성

이를 통하여 충분한 episode를 시행시키면 무작위 정규분포의 데이터들이 각 시점에 발생하여 데이터로 입력되게 되는 것이다. <그림 31>의 두 번째는 같은 방법을 통하여 63일의 Simulation 데이터를 생성했을 경우 역시 무작위 정규분포형을 나타내고 있음을 보여주고 있다.

나. DQN 구현

a. 환경설정

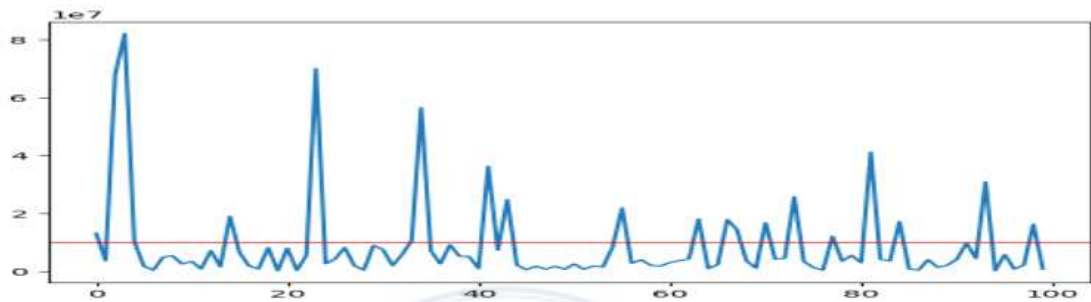
선택된 10개의 종목 데이터 중 Simulation 시행으로 비교 가능 길이 (6,000일 이상) 및 백 테스트를 위한 변동성 고려한 3개 종목의 데이터를 사용하여 test를 위한 데이터(2,000일)를 나눠두고, 나머지 데이터를 Train 데이터로 사용하였다. 각 Step의 상황(state)은 일일(day)로 설정한다. 각 일의 포트폴리오 구성비는 DQN에 의하여 결정되고 상황(state)에 따른 보상은 일일 수익률로 계산되어 진다. 각 보유는 그날 현금화하여 평가하고 이를 비중으로 재구매했을 경우 각 종목의 비중 구매 후 남은 현금은 현금보유 한다. 거래수수료 및 현금화 자체의 이자율은 계산하지 않는다. 기초 투자금은 일천만 원으로 설정하였다.

b. 훈련과정

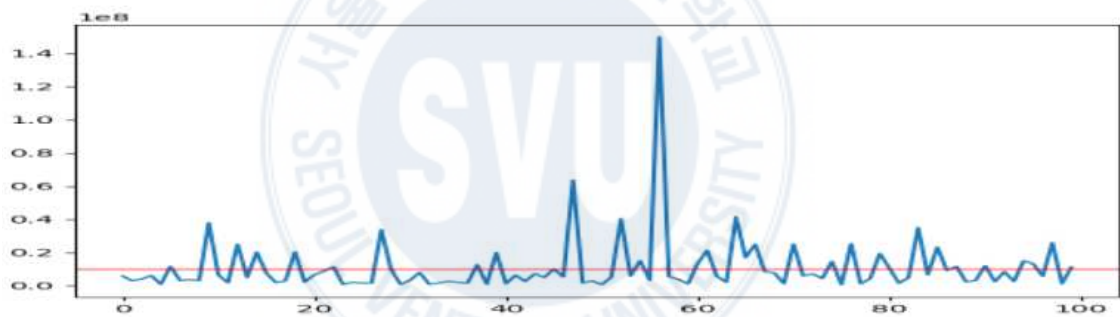
DQN 훈련에는 전체 Train 데이터를 이용한 학습, Simulation을 사용하여 생성하는 2가지 형식의 데이터 집합(63일 window 무작위 선택 및 GBM Monte Carlo Simulation 생성 63일과 252일)을 이용한 학습 총 3가지를 실시하였다. 먼저, 전체 데이터를 이용한 Train 과정의 결과를 보여주는 [그림 3-30]이다. 1000회(episode)의 훈련과정 중 일정 episode 시점에서의 학습된 가중치로 검증 100회 실시한 결과이다. 결과 훈련 episode가 증가함에 따라 최대 금액과 평균 결과 금액이 증가하는 것을

확인할 수 있다. 가로 실선은 일천만 원 초기 투자 금액이다.

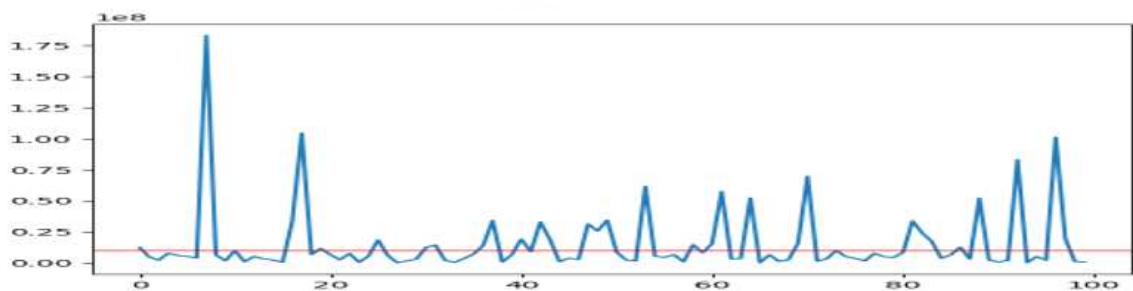
train_episode = 100, result_mean = 8903278.42, try = 100



train_episode = 500, result_mean = 10941969.98, try = 100



train_episode = 700, result_mean = 15215491.33, try = 100

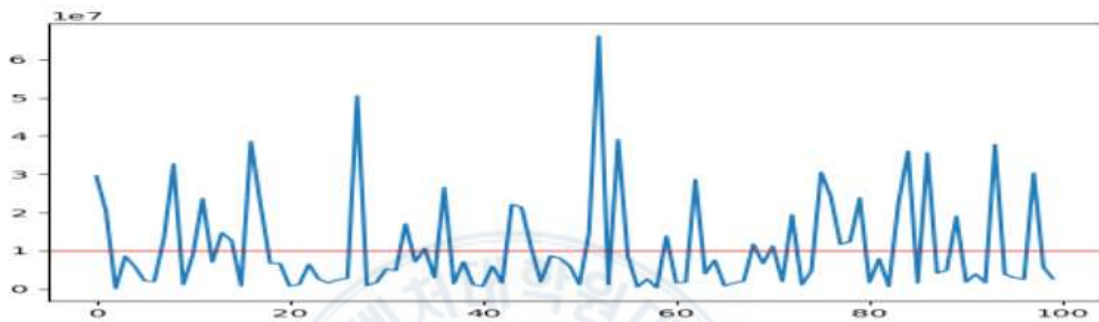


[그림 3-31] 전체 데이터 훈련의 훈련과정 확인 투자 100회

다음은 GBM MC Simulation을 통해 생성한 63일 데이터를 이용하

여 DQN을 훈련하였다. 훈련데이터의 episode 길이가 짧은 것을 참작하여 훈련 episode 횟수를 증가하여 10,000회 실시하였다.

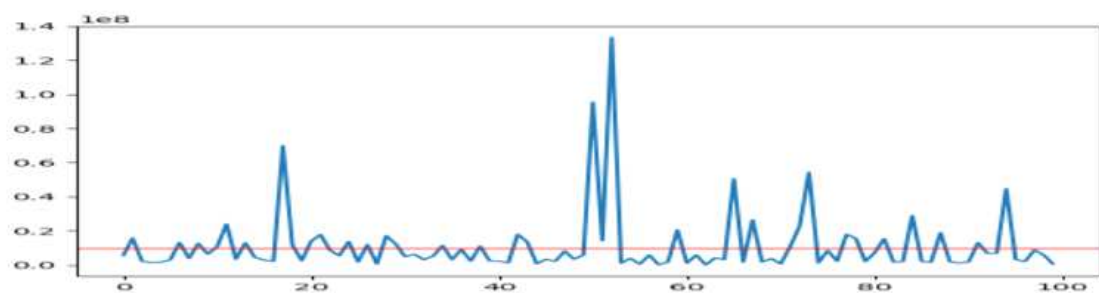
train_episode = 2000, result_mean = 10819688.53, try = 100



train_episode = 7000, result_mean = 11870377.74, try = 100



train_episode = 9000, result_mean = 11245558.95, try = 100



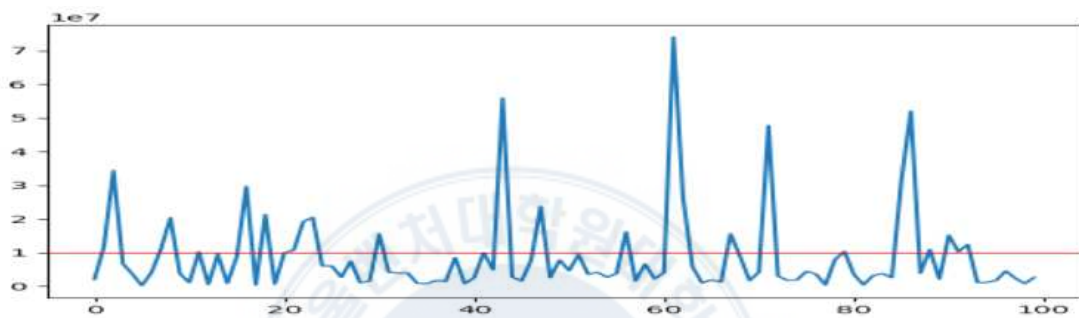
[그림 3-32] 63일 Simulation 데이터 훈련과정 확인 투자 100회

Simulation을 통한 데이터에서는 최대 금액의 증가는 크지 않아 보인다.

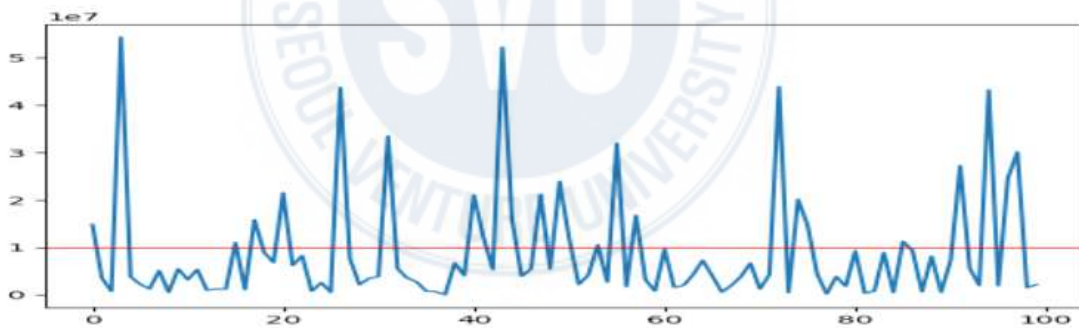
그러나 기준선(초기 투자금 1천만 원)에 이상으로 전체적인 그래프가 접근하는 모습은 전체 Train 데이터를 사용한 것과 유사하다.

세 번째는 Simulation을 이용하여 252의 train 데이터를 생성하는 경우이다. 총 3500 episode를 학습시켰다.

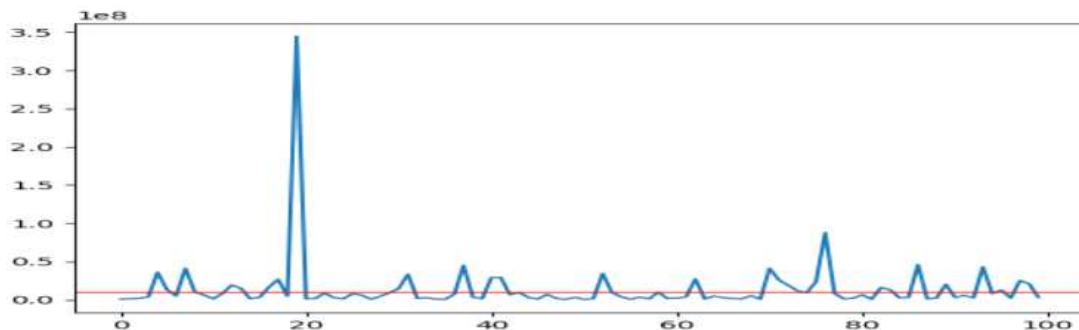
train_episode = 1000, result_mean = 8657784.31, try = 100



train_episode = 2500, result_mean = 8938321.16, try = 100



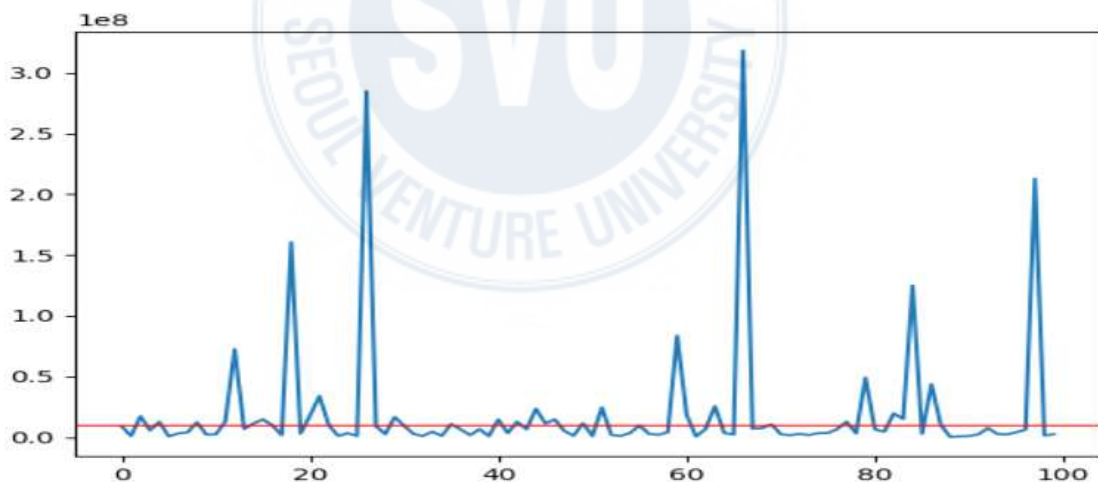
train_episode = 3500, result_mean = 13798265.03, try = 100



[그림 3-33] 252일 Simulation 데이터 훈련과정 확인 투자 100회

252일을 생성하여 훈련의 경우 평균이 훈련에 따라 증가하였다. 그러나 훈련 중간 특정 episode에서 이상치(극단적인 금액 차이) 수익상승이 발생하여 episode에 의해 평균이 높게 나타나는 경우가 있었으나, 안정적인 결과라고 볼 수 없고 데이터에 맞는 가중치 정책이 낮은 확률로 훈련 데이터의 변화량에 맞게 나타난 것으로 보인다. 예를 들어 [그림 3-33]과 같이 1500 episode를 훈련한 가중치로 Test 하였을 경우 평균이 19759426.8로 높게 나왔으나 훈련이 지속 되면서 오히려 2500 episode에서 낮아짐을 볼 수 있다. 이는 simulation 된 데이터가 Test 데이터에 좋은 결과가 나올 수 있는 비중으로 작은 확률로 일치했을 가능성이 있다. 그러나 장기적인 학습에서 안정적인 결과로 볼 수 없다.

train_episode = 1500, result_mean = 19759426.8, try = 100



[그림 3-34] 252일 Simulation 데이터 1500회 훈련 확인 투자 100회

c. 결과분석

상황(state)에 대한 평가에 사용된 신경망이 가장 기본적이고 작은 MLP(2 layer-multi layer perceptron)을 사용하였으나, DQN 모델이 훈련

이 진행됨에 따라 점점 안정적이고 좋은 수익률을 획득할 확률이 높아지고 있다. 처음에는 검증 100회 중 일천만 원 이하의 수익률 결과가 다수 보이지만 훈련이 진행되면서 모든 테스트의 경우 일천만 원 선에 접근하거나 높아지는 것을 볼 수 있다. 이는 DQN 모델이 훈련을 통하여 10개 종목의 주가 상황에 따른 정책을 학습함을 알 수 있다.

다음으로 3종류의 데이터에 대해서 특히, GBM Monte Carlo simulation을 통하여 생성된 데이터에 대해서도 학습이 이뤄지고 있음을 알 수 있다. 3000일 또는 그 이상의 가상 Simulation 데이터에 대한 훈련과 충분한 훈련횟수의 결과 데이터가 더 필요해 보이나 현재의 실험으로도 가상의 Simulation을 통한 데이터를 통해 강화학습 알고리즘이 상황에 맞는 포트폴리오 구성을 학습함을 알 수 있었다. 데이터의 길이가 길수록, 훈련의 회수가 늘어날수록 높고 안정적인 결과를 보임을 알 수 있다. 따라서, KOSDAQ 시장과 같은 2~3년의 짧은 데이터에도 강화학습 알고리즘을 활용할 가능성이 있다.

즉, 상장 기간이 짧아 기존의 데이터를 이용하여 포트폴리오를 구성할 때 앞으로의 변동성에 대한 예측이 곤란할 경우 가상의 Simulation 데이터를 통하여 포트폴리오 강화학습 모델을 훈련 시켜 대응할 가능성이 있다. 기존의 모델은 과거의 데이터를 이용한 현재의 상황과약이나 앞으로의 상황변화를 예측하는 알고리즘이었다. 이러한 분석을 통하여 포트폴리오 비중 설정은 별도의 전략으로 수립되었다 볼 수 있다. 그러나 강화학습을 이용한 포트폴리오 구성은 상황 평가에서 포트폴리오 구성비의 선택까지 일련의 과정을 한꺼번에 학습한다고 볼 수 있다.

특정한 상황에 대한 분석보다 정책의 결과에 따라 모델의 학습이 이뤄진다. 그러므로 데이터로부터 발생하는 정보가 실제 상황이 아니더라도

정책 학습이 가능하다. 단, 이러한 정보로부터 학습한 결과가 꼭 좋은 성과를 보장하지 않는다.

특히, DQN은 학습한 정보에 따라 결과 변동 폭이 큰 것으로 알려져 있다. 좋은 가상의 데이터를 다시 말해서 좋은 강화학습 환경(Environment)을 설정하는 것이 중요하다. 환경설정이 구체적이면 그로부터 가상 Simulation 데이터도 생성할 수 있다.

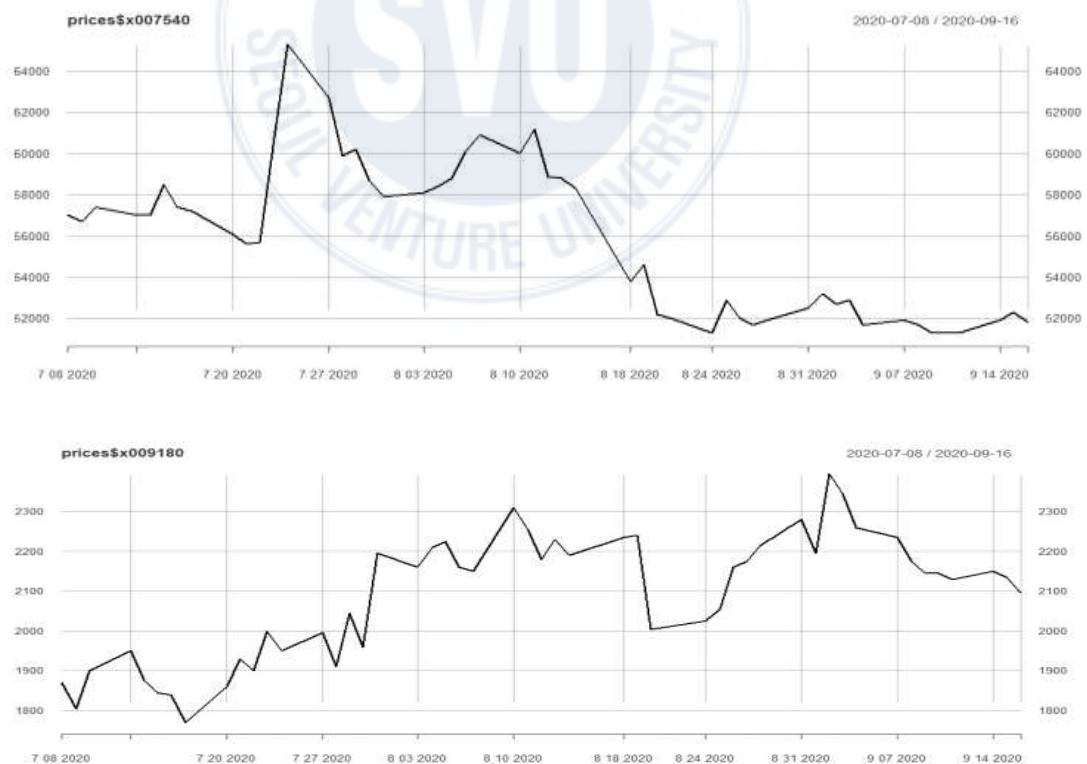


Ⅳ. 백 테스트(Back Test) 및 분석

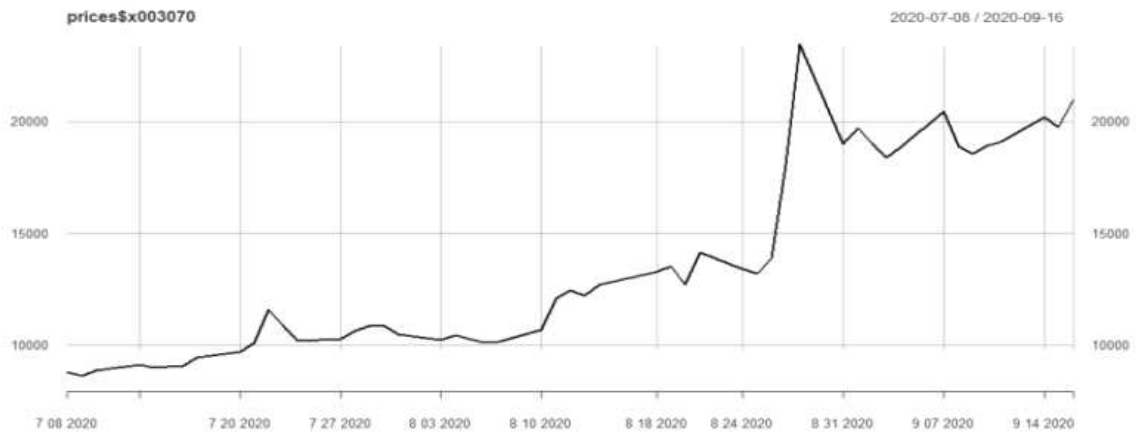
1. 백 테스트 환경설정

가. 데이터의 전처리

백 테스트 비교 분석을 위하여 종목 선택에서 선정된 10개 종목의 2020년 7월 9일에서부터 2020년 9월 16일까지의 데이터를 추가로 수집하고 이 중 3종목에 대하여 포트폴리오 구성 방법의 데이터로 사용하였다. 선정 종목은 코오롱 글로벌(X003070), 샘표(X007540), 한솔로지스틱스(X009180) 이다.



[그림 3-35] 백 테스트에 사용된 3종목의 주가



[그림 3-35] 백 테스트에 사용된 3종목의 주가

나. 환경설정

일일 수익을 기준으로 포트폴리오 동적 비중 계산 시는 10일 즉 2주간의 수익률을 보고 포트폴리오 비중을 재설정한다. GARCH와 RNN의 경우는 하루 앞의 예측치를 이용하여 같은 방법을 적용한다. 강화학습의 경우 매일 평가하여 현금화하는 것으로 설정한다. 각각의 각 포트폴리오 전략에 따라 매매수하고 남은 현금은 현금으로 유지 평가 시 합산하여 계산하였다. 단, 거래 시 발생하는 수수료와 비용은 없는 것으로 설정하고 제외하였다.

2. 테스트 결과 및 분석

가. 최소변동성 기준 동적 비중 전략

10일 즉, 과거 9일과 오늘(종가:close)의 수익률로 매일 비중을 동적으로 최소변동성으로 계산하여 매매하는 전략으로 테스트 기간의 주식을

거래한 결과는 다음 [그림 3-35]이다.



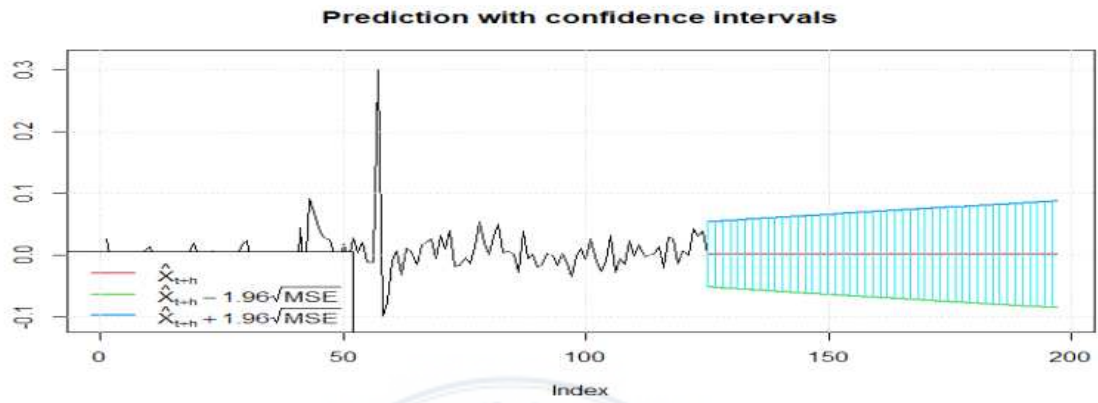
[그림 3-36] 최소변동성 기준 동적 비중 전략

8월 0.2(20%) 가까이 수익률을 발생하였으나, 곧 떨어져 수익률이 거의 없는 상황까지 갔다. 이시기는 선정한 종목의 가격 변화 그래프에서 2종목이 가격이 감소하는 시기와 유사하다. 그리고 다시 상승하여 최종 0.125(12.5%)에 머무르고 있다. 누적수익률이 진폭이 발생하고 있다.

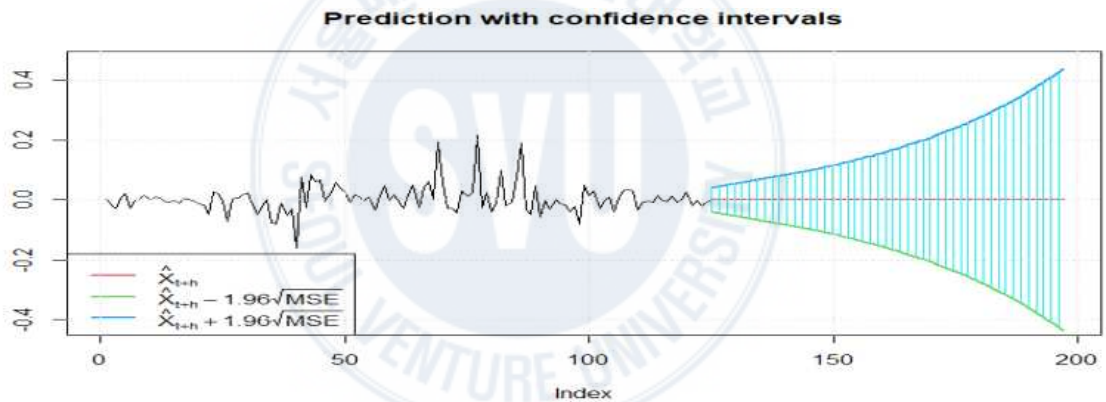
나. GARCH를 이용한 변동성 예측

지난 500일의 데이터를 이용하여 각 종목의 GARCH 모형 및 변동성을 예측하였다.

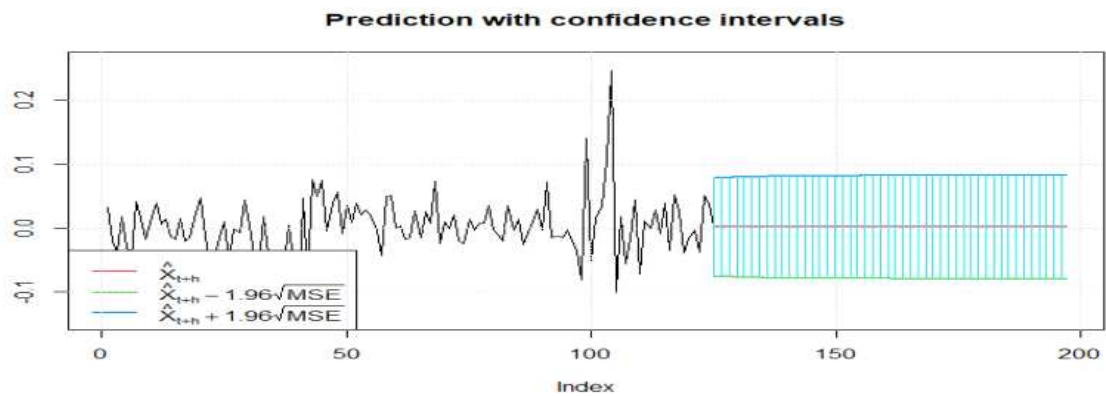
x003070 종목의 향후 73일에 volatility 대한 예측



X007540 종목의 향후 73일에 volatility 대한 예측



X009180 종목의 향후 73일에 volatility 대한 예측

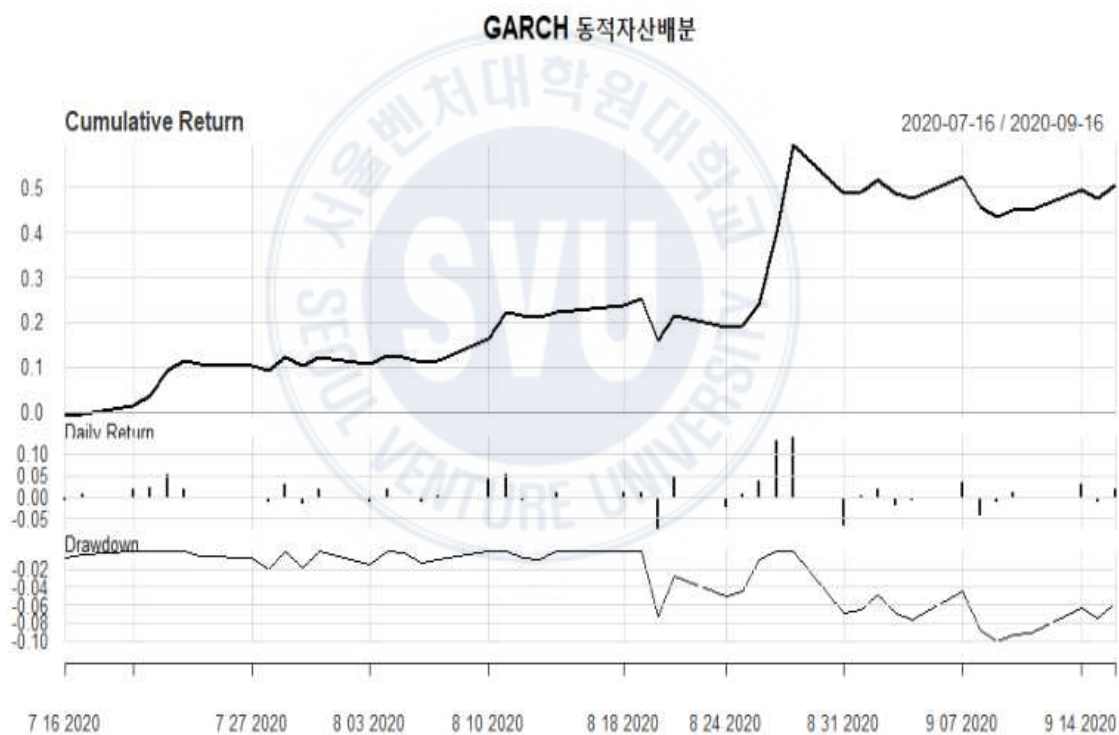


코오롱 글로벌(X003070), 샘표(X007540), 한솔 로지스틱스(X009180)

[그림 3-37] GARCH를 이용한 변동성 예측

각 종목의 모형으로부터 예측 표준편차를 구하여 공분산을 예측하였다. 이를 앞의 최소변동성 기준 동적 비중 전략(10일 간격)을 사용하였다.

최소, 최대 소유 비중 한계를 설정하지 않아 변동성이 가장 작은 종목에 대하여 최대로 유지하는 경향을 보이기 때문에 예측 GARCH 모형에서 변동선 예측이 가장 작은 종목의 변화 그래프와 유사한 것을 볼 수 있다.



[그림 3-8] GARCH를 이용한 최소변동성 기준 동적 비중 전략

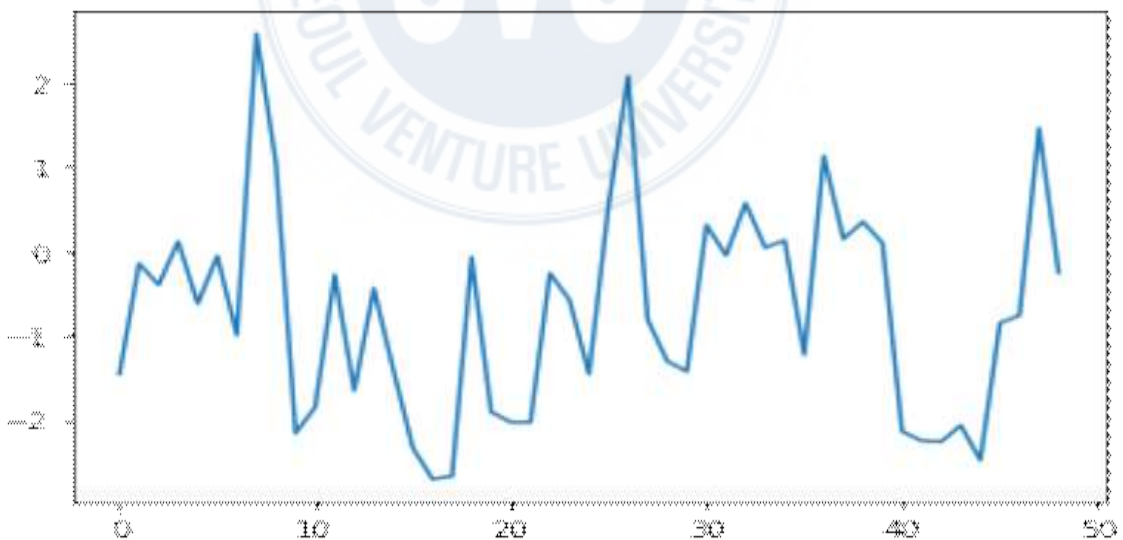
전체적으로 안정적인 수익률을 발생시키고 있다. 상승 특히 8월 24일 이후 상승 국면은 3종목 중 가격이 상승한 종목의 비중을 잘 선택하여 유지하였다고 볼 수 있다. 이는 GARCH 모형을 통해 예측한 변동성이 적은 것을 선택했음을 알 수 있다. 그 결과 x003070과 x009180의 가격 상승을

취할 수 있었다. 결과적으로 최종 0.5(50%)에 이르고 있다. 최소변동성 기준 동적 비중 전략에서 누적수익률에 변동 진폭이 있던 것과는 달리 누적 수익률의 상승이나 평탄함을 유지하는 것을 볼 수 있다.

다. RNN을 이용한 수익 변동 예측

RNN을 이용하여 전 10일간의 주가를 로그 차분한 데이터를 입력으로 앞의 1일의 예측 로그 변화를 구하였을 때 다음의 [그림 3-38]은 매번 슬라이딩 전진 된 일일의 로그 수익률을 49일간 예측한 것이다. 이러한 예측치를 가지고 과거 10일의 백 테스트 기간의 데이터와 하루 앞의 1일의 예측치를 이용하여 변동성을 계산하여, 10일마다 최저 변동성

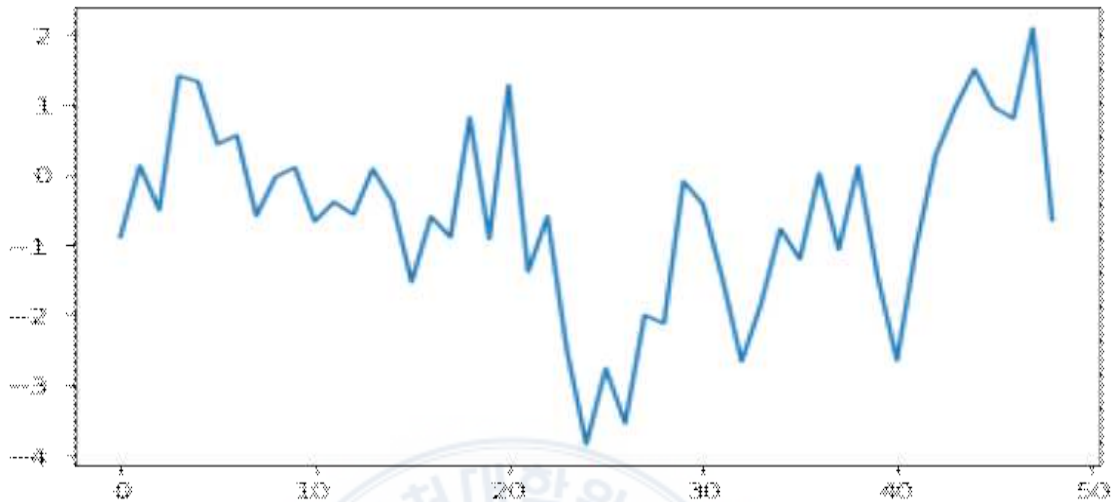
[x003070] : look back = 10 days, predict = after 1day



코오롱 글로벌(X003070), 샘표(X007540), 한솔 로지스틱스(X009180)

[그림 3-39] RNN을 이용한 수익 변동 예측

[x007540] : look back = 10 days, predict = after 1day



[x009180] : look back = 10 days, predict = after 1day



코오롱 글로벌(X003070), 샘표(X007540), 한솔 로지스틱스(X009180)

[그림 3-39] RNN을 이용한 수익 변동 예측

비중을 새롭게 계산하는 동적 비중 전략을 사용하여 테스트를 진행하였고, 다음 [그림 3-39]은 그 결과이다.

RNN 예측 동적자산배분



[그림 3-40] RNN을 이용한 최소변동성 기준 동적 비중 전략

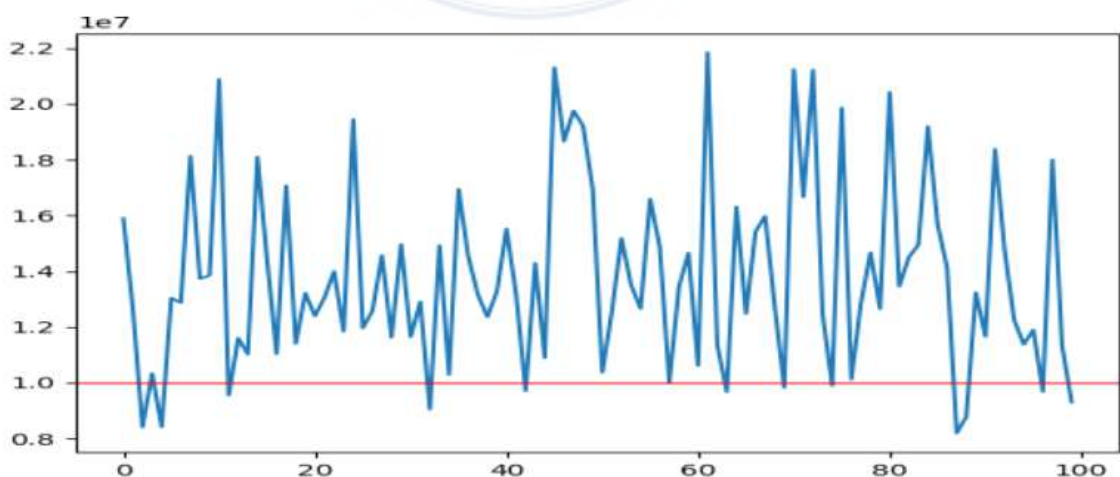
전체적으로 안정적인 수익률을 발생시키고 있다. GARCH를 이용한 경우와 같이 8월 24일 이후 상승 국면에서 3종목 중 가격이 상승한 종목의 비중을 잘 선택하여 유지하였다고 볼 수 있다. 최종 0.3(30%)에 약간 미치지 못하고 있다. RNN이 훈련과정에서 수익률의 변동성이 GARCH 모델과 같이 반영되었음을 알 수 있다. GARCH를 이용한 모델과 다르게 누적수익률이 부드럽게 상승하고 있다. 이러한 점을 반영하듯 낙폭에 대한 부분을 보면 GARCH를 이용했을 때 보다 안정적임을 알 수 있다. 본 연구의 백 테스트가 단기적인 누적수익률의 결과라는 점을 고려할 때, GARCH 모델을 이용한 누적수익률이 높게 RNN을 이용한 예측모델보다 누적수익률이 높게 나타났으나 장기적인 안정성에 대하여는 본연구로는 알 수가 없어 보인다.

라. DQN 및 GBM-MC-DQN

아. DQN

먼저, 전체 데이터를 사용하여 훈련한 가중치를 사용하여 백 테스트 기간의 일일 거래 전략에 대하여 알고리즘 거래를 100회 실시하였다. [그림 3-39] DQN 강화학습 포트폴리오 투자 100회의 결과를 보면 손실의 경우 -20%에 가까운 손실도 발생하였으나 전체적으로 수익이 발생함을 알 수 있다. 최대 이천일백만 원(110%)에 달하는 누적수익률을 보이기도 한다. 100회에 대한 평균을 내어 기댓값을 살펴보면, 약 천삼백 팔십만 원(38%)의 수익률 평균을 보인다. 즉, DQN 알고리즘으로 일천만 원을 투자했을 때 평균 기대할 수 있는 수익률은 38%이다. 이는 GARCH 모델을 이용한 경우보다 낮고 RNN을 이용했을 때보다는 높다. 그러나 아직 추가 학습이 가능하고 최대수익률이 높은 점을 고려한다면 좀 더 위험하고 좀 더 높은 수익률을 내는 전략이라 할 수 있다.

weight = base, return_mean = 13834357.4, try = 100

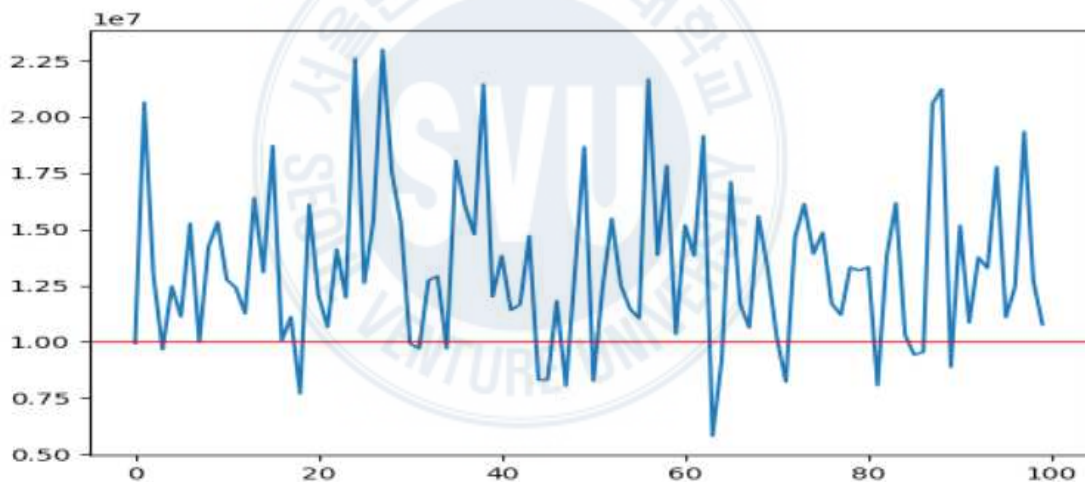


[그림 3-41] 전체 데이터 DQN 강화학습 포트폴리오 테스트 투자 100회

b. GBM-MC-DQN (63days)

다음 [그림 3-41]은 Simulation을 통해 생성한 데이터로 훈련한 가중치를 사용하여 백 테스트 기간의 거래를 100회 실시한 결과이다. 전체 데이터를 사용한 결과보다 약간 더 손실의 경우가 있기는 하지만 전체적으로 크게 다르지 않음을 알 수 있다. 최대 이천이백오십만 원(125%) 수익률을 발생시키는 반면 -40%에 달하는 원금 손실의 경우도 발생하고 있다. 수익률의 평균은 33.5%이다.

weight = 63days_sim_data, return_mean = 13354326.85, try = 100

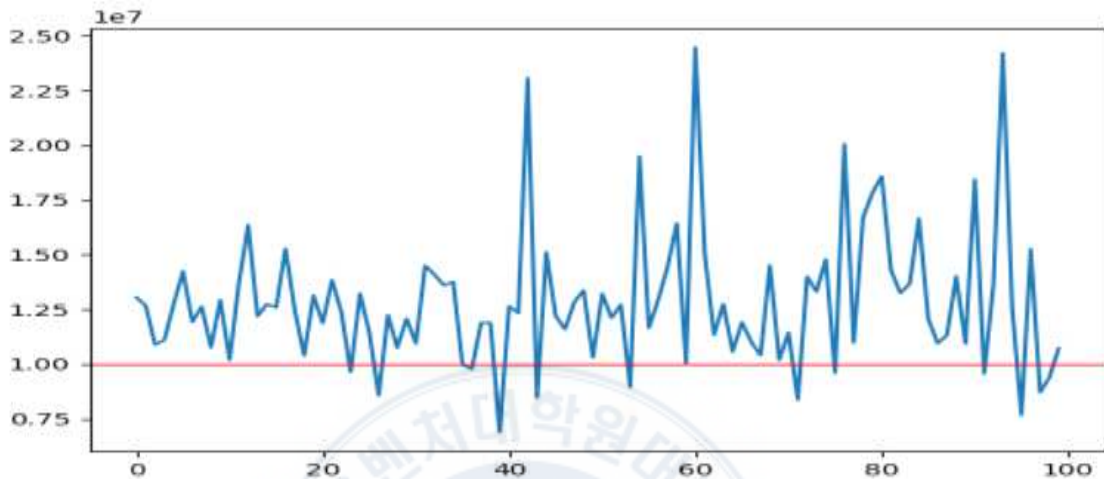


[그림 3-42] GBM-MC(63일) DQN 강화학습 포트폴리오 테스트 투자 100회

c. GBM-MC-DQN (252days)

252일의 Simulation 데이터를 이용하여 학습시킨 가중치로 테스트를 진행한 결과 최대수익률은 이천사백만 원(140%)으로 높아졌다. 그러나 역시 손실의 경우가 발생하는 것을 볼 수 있다. 평균 수익률은 28.4%임을 알 수 있다.

weight = 252days_sim_data, return_mean = 12841179.1, try = 100



[그림 3-43] GBM-MC(252일) DQN 강화학습 포트폴리오 테스트 투자 100회

d. 데이터 종류 간 비교

GBM-MC로 생성한 Simulation 데이터로 학습시킨 가중치와 전체 데이터로 학습시킨 데이터 모두 학습이 가능한 것으로 나타났다. 그러나 Simulation 데이터의 경우 백 테스트 결과의 누적수익률의 변동 폭이 더 큰 것으로 나타났다. 이 같은 현상은 Simulation 데이터를 더 긴 기간으로 생성할수록 누적수익률 변동 폭이 커지는 것으로 나타났다. 이는 생성된 데이터의 무작위 정규분포의 범위도 커지기 때문으로 보인다. 그만큼 더 많은 환경(Environment) 모수의 경우를 학습하여야 하기 때문이다. 최적 행동(Action)을 결정하는 최적 정책(Policy)에 도달하는데 더 많은 학습이 필요해 보인다. 이는 보상(Reward)의 누적인 결과(Return)로 최대누적수익률 크기도 커지며 반대로 최소누적수익률 또는 손해의 크기도 커지는 불안정한 결과로 나타나고 있다.

결국, Simulation 데이터로 학습하는 모수 환경에서 발생 가능한 경우의 수가 더 많다는 것을 의미하며, 그에 따라 정책의 최대성과 결과의 변동 폭도 커지는 것이다. 이는 투자결과에 대한 안정성과 관련이 있다. 성과에 대한 기댓값을 예측하는데 어려움을 줄 수 있다. 안정성 측면에서는 위험 즉, 변동성이 높다고 할 수 있다. Simulation 데이터를 이용한 각 경우의 100회의 누적수익률 평균을 통해서 기대수익률을 살펴보면 오히려 조금씩 낮아지고 있다.

이러한 현상이 Simulation 데이터 사용 길이와 관련 있는지 또는 학습 반복 횟수와 관계가 있는지는 각각의 데이터 길이와 학습 횟수를 더 다양하게 변경하며 비교하는 추가 실험 분석이 필요해 보이며, 그러한 실험 결과를 투자 횟수 100회가 아닌 2000회 이상 증가시켜 일반화하는 추가 연구가 필요하다.

V. 결 론

1. 연구결과 요약 및 시사점

본 연구는 주식 데이터를 분석하여 투자지표를 설정하고 그를 통해 종목을 선정하였다.

이렇게 선정된 주식시장 종목을 첫째, 기초 통계분석으로 공분산을 구하여 포트폴리오를 구성하는 방법을 구현하였다.

둘째, 변동성(Volatility)을 예측하는 GARCH 모델을 구성하여, 변동성을 예측하고 이를 통하여 포트폴리오를 구성하였다.

셋째, Deep Learning 방법 중 연속적인 데이터 분석에 사용하는 RNN 모델을 이용한 예측치로 포트폴리오를 구성하였다. 마지막으로 DQN 강화학습을 이용하여 주가 상황 분석에서 포트폴리오를 구성까지 일괄적으로 학습하는 모델을 구현하였다.

이때, DQN을 이용한 포트폴리오 구성 및 투자 방법은 전체 데이터를 이용한 학습, GBM-MC Simulation을 이용하여 2가지 종류의 데이터를 생성하여 학습하는 총 3가지 훈련을 진행하였다. 이렇게 구현한 방법들의 결과를 비교하기 위하여 2개월간의 데이터를 다시 확보하여 백 테스트를 진행하고 결과를 비교 분석하였다.

이러한 분석 실험과정을 통하여 살펴본 결과는 기초 통계를 통하여 구한 공분산으로 포트폴리오를 구성한 경우 과거부터 현재까지의 정보를 통해 분석한 변동성으로 구성되어있어 새로운 주가 변화에 대응하지 못하고 누적수익률이 낮아지는 현상을 보였으며 그러한 변화가 정보로 반영된 후

에야 다시 누적수익률이 발생함을 볼 수 있었다. 다음으로 GARCH를 통해 변동성(Volatility)을 예측하는 방법으로 포트폴리오를 구성하면 최소변동성 포트폴리오 구성 전략에 따라서 변동성이 적은 종목으로 투자가 집중됨을 볼 수 있다. 이는 주가 변화예측에 따라 포트폴리오 구성이 대응함을 보여준다.

RNN을 예측치로 공분산을 구해 포트폴리오를 구성한 경우는 GARCH와 같이 주가 변화에 대응하는 것을 알 수 있다. 그러나 누적수익률 결과는 GARCH보다 성과가 낮아 보였다. 하지만 특정 수익률 고점에서 저점(Draw Down) 그래프를 보면 GARCH 모형과 RNN 모형 모두 급격한 변화가 기초통계계산을 통한 투자에 비하여 적은 형태를 보이지만 RNN의 경우가 GARCH보다 안정적인 모습을 확인할 수 있었다. 이 같은 결과가 일반적인 결과라면 장기적인 투자에서 RNN 분석을 통한 포트폴리오 구성이 더욱 안정적인 결과를 내놓을 가능성이 있다.

DQN 강화학습을 이용한 투자는 훈련된 가중치를 이용하여 같은 조건에서 100회 시행한 누적수익률을 통해 살펴보았다. 이때 100회 중 손해가 발생하기도 하였으나 반대로 GARCH 또는 RNN 분석을 이용한 방법보다 높은 누적수익률을 보이기도 하였다. 100회 누적수익률에 대한 평균을 통해서 기댓값을 추정하면 기존의 방법과 누적수익률이 유사함을 알 수 있었다. 다만 더 높은 누적수익률 가능성이 있는 반면에 손실의 위험도 존재함은 포트폴리오 전략의 기본 목적인 위험분산과 반대되는 결과이다.

DQN 강화학습 방법을 Simulation 데이터를 이용하여 학습한 가중치로 백 테스트를 진행한 결과가 실제 데이터를 이용한 누적수익률 결과보다 최고누적수익률과 최고누적손실률 간의 폭이 더 큼을 알 수 있다. 즉, 최고성과는 증가하는 반면 안정성은 더 떨어질 수 있음을 의미한다. 이러한

현상은 63일과 252일 Simulation 데이터로 훈련한 가중치의 백 테스트 결과를 비교해 볼 때 Simulation으로 생성한 데이터의 길이가 길어질수록 증가한 것으로 나타났다. 이는 포트폴리오 구성에 대한 가능성 있는 환경에 대하여 더 넓은 범위의 모수에 대하여 학습한 결과라고 생각할 수 있다. 더 많은 상황에 대한 학습은 실제 투자에서 더 좋은 결과를 가져올 수도 있으나 충분한 학습이 이뤄지지 않는다면 안정적 결과를 기대할 수 없음을 나타낸다.

강화학습인 DQN 학습이 실제 데이터가 아닌 가상의 무작위 생성 데이터를 통하여 실제 데이터와 유사한 결과를 보인다는 것은 상당히 고무적이라고 할 수 있다. DQN 강화학습을 사용한 포트폴리오 투자의 안정성은 내부의 학습 네트워크가 가중치가 힘을 잃게 되는 기울기 소실(Gradient Vanishing)이 일어나지 않는 한 최대한 학습을 통해 높일 수 있을 것이다. 가장 기본적인 MLP로 구성된 네트워크를 개량하는 방법도 고려할 수 있다.

결과적으로 DQN 강화학습을 이용한 포트폴리오 투자는 아직 손실의 위험성은 있으나 충분히 기대 누적수익률을 낼 가능성이 있으며, 이러한 모형의 학습은 실제 데이터가 아닌 Simulation 데이터를 이용하여도 충분히 가능하다는 것을 알 수 있다.

2. 연구한계 및 향후 연구 방향

본 연구의 결과를 통해 실제 데이터가 아닌 Simulation 데이터를 이용하여 DQN 강화학습을 이용한 포트폴리오 투자 모형의 학습이 가능함을 알 수 있었다. 그러나 Simulation으로 생성한 데이터의 길이가 길어질

수록 최고누적수익률과 최고누적손실률 간의 폭이 더 커져 안정성이 낮아진다. 이는 포트폴리오 전략의 위험분산이라는 목적과 반대되는 결과라고 할 수 있다.

이러한 문제는 학습 횟수를 증가시켜 가능한 환경(Environment) 모수의 경우의 수를 더 많이 학습시켜 해결할 수 있을 것이다. Simulation을 이용한 학습이 전체 원본데이터를 학습하는 경우와 다른 점 중 하나가 episode의 길이를 설정할 수 있다는 점이다. 짧게 episode를 설정할 경우 몇 배 빠르게 학습할 수 있으며 그에 따라 더 많은 학습 횟수를 같은 시간에 할 수 있다. 반대로 실제 원본데이터보다 긴 데이터로 학습할 경우는 비례적으로 더 많은 학습을 통해서만 안정적인 누적수익률 상승효과를 가져올 것이다. 그러나 지속적인 학습은 심층 네트워크의 가중치가 힘을 잃는 Gradient Vanishing 한계가 있다.

이러한 문제에 대하여 첫째, 강화학습 Simulation 데이터 길이와 누적수익률에 대한 관계를 추가적인 연구가 필요해 보인다. 가상의 환경(Environment) 설정이 잘 이뤄졌는지는 실제 데이터를 통한 반복된 백-테스트 결과를 누적 분석하여 알 수 있다. 이를 통해서 Simulation을 이용한 데이터 길이에 대하여 비용과 결과 대비 최적으로 효율적인 길이를 구하는 방법 연구가 필요하다.

둘째, Simulation 데이터를 이용한 3가지 DQN 강화학습의 경우 100회의 누적수익률 평균을 통해서 기대수익률을 살펴보면 오히려 조금씩 낮아지고 있다. 이러한 현상이 Simulation 데이터 사용 길이와 관련 있는지 또는 학습 반복 횟수와 관계가 있는지는 각각의 데이터 길이와 학습 횟수를 더 다양하게 변경하며 비교하는 추가 실험 분석이 필요해 보이며, 그러한 실험 결과를 투자 횟수 100회가 아닌 2000회 이상 증가시켜 일반화

하는 추가 연구가 필요하다.

셋째, 여러 강화학습 방법의 적용 또는 개선 대한 연구 및 정규화(regularization), Drop out 등의 내부 심층 네트워크의 개량을 통한 기울기 소실(Gradient Vanishing)을 방지에 관한 연구가 필요하다.

본 연구는 DQN 강화학습을 이용한 포트폴리오 구성비율 학습 과정을 이해하고자 금융 시계열 분석방법을 개괄해 살펴보았다. 그러나 실제 금융투자전략으로서 어느 방법이 우세한지는 개괄적인 본연구의 시현 방법론만으로는 부족하다. 각 방법의 장점을 살피거나 세부적인 전략과 개량전략을 추가함으로써 결과는 달라지기 때문이다.

더구나 강화학습과 Deep Learning Network는 지금도 급속도로 발전하고 있는 방법으로 본연구에서 사용한 DQN 이외에도 정책을 위주의 Policy Optimization 강화학습 계열과 본연구에서 사용한 Dynamic Programming 계열이 있으며, 두 계열을 융합한 Actor-Critic 방법 계열도 존재한다. 따라서 본 연구에서 사용한 가장 간단한 DQN 강화학습 이외의 여러 기법의 강화학습을 적용해 볼 필요가 있다. 같은 강화학습 방법이라 하여도 내부의 학습 네트워크나 행위 선택 방법 등을 개량할 수도 있다.

앞에서 살펴본 기존 예측 방법론들도 강화학습 방법에 적용하여 발전시킬 수 있다. 실제로 로봇공학이나 Ai 연구에서는 정책을 결정하는 과정에 RNN을 접목하거나 환경 상황파악에 CNN과 같은 방법도 사용되고 있다. 내부의 네트워크 강화를 위하여 여러 통계적 예측 기법들이 적용되어 사용되고 있다.

이러한 방법론이 실제 경영이나 투자에 사용되기 위해서는 단순히 기

기계학습 모델을 이용한 투자성과만을 보고 실효성을 판단하기에 위험성이 크다. AI 학습환경 설정 조건이 항상 유효하지는 않기 때문이다. 현재 학습시킨 모델의 학습환경 설정의 유효성을 판단하기 위해서는 투자 행위를 결정하는 기계학습의 작동방식에 대한 이해가 필요하다. 기존 기계학습 논리적 작동과정에 대한 이해를 투자 측면에서 재해석하는 연구가 필요하다. 그러한 연구를 통해서 기계학습 방법을 투자전략에 맞게 개선해가는 연구가 가능해질 수 있다.

더해서 이러한 알고리즘을 통한 금융투자 분야뿐만 아니라 모든 SMART 분야에 해당하는 문제로서, 이러한 AI 적용을 연구하는 모든 분야에서 어떻게 현상을 정형화한 환경(Environment)으로 조작 정의하느냐는 중요한 화두이다.

참 고 문 헌

<국내 문헌>

- 강문주(2018), 회귀 강화학습을 이용한 포트폴리오 자산 배분의 성능 향상 연구, 인하대학교 대학원 컴퓨터공학과, 석사학위논문
- 강장구(KAIST), 류두진 (KAIST)(2009). “옵션시장에서 GARCH계열 모형들의 성과 비교에 관한 연구”, 한국증권학회지 제38권 2호 137-176
- 김경목(2017). “머신러닝과 수급분석을 활용한 주식 포트폴리오 구성 연구”, 국민대학교 비즈니스IT 전문대학원, 석사학위논문
- 김경진(2020). ”GARCH와 LSTM 모형을 이용한 수입육 가격 변동성 예측력 비교“, 건국대학교 대학원, 석사학위논문
- 김기범(2018), 수리적 모델과 딥러닝을 이용한 코스피200 선물변화 예측, 아주대학교 일반대학원, 석사학위논문
- 김상용·이진아(2009). “일반 자기회귀 이분산 모형을 이용한 시계열 자료 분석”, 한국데이터정보과학회지, 2009, 20(3) p475-483
- 김주봉(2019), 비동기 심층 강화학습 기반 금융 포트폴리오 관리, 한국기술교육대학교 대학원 컴퓨터공학전공 석사학위논문
- 김주봉·허주성·임현교·권도형·한연희(2018). “A3C를 활용한 블록체인 기반 금융 자산 포트폴리오 관리”, KIPS Trans. Comp. and Comm. Sys. Vol.8, No.1 pp.17~28 pISSN: 2287-5891
- 김진석(2020), “강화학습을 이용한 모바일 로봇의 국지적 동적 장애물 회피, 과학기술연합대학원대학교, 생산기술(로봇공학)전공,

박사학위논문

- 박정빈(2019), RNN을 이용한 금융 시계열의 고저가 수익률 예측 연구, 숙명여자대학교 대학원, 석사학위논문
- 박한상(2020), 인공지능 기술을 활용한 금융시장 분석과 포트폴리오 최적화, 가천대학교 일반대학원, 석사학위 논문
- 이종원(2018). 딥러닝 모델 최적화 기반 순차 데이터 예측 시스템, 배재대학교 대학원, 박사학위논문
- 이현열(2019), R을 이용한 퀀트 투자 포트폴리오 만들기, 제이펍, Ch 9, Ch 11, p195 - p256
- 오병훈(2018), 딥러닝을 이용한 무기체계 수리 부속의 간헐적 수요예측, 고려대학교 컴퓨터정보통신대학원, 석사학위논문
- 임준범(2019), 확장된 이산 행동 영역에 적용한 강화학습 기반의 행동 특화된 전문가 모델 앙상블 트레이딩 시스템, 아주대학교 대학원 금융공학과, 석사학위논문
- 장수영 외(2019), “심층 강화학습 기술 동향,” 전자통신동향분석 34권 제4호, 2019. 8, pp. 1-14.
- 정용진(2019). “순환신경망 기반 비선형 ARMA-GARCH 모형을 이용한 S&P500 지수 예측”, 서울대학교 대학원, 석사학위논문
- 조성근(2018), 암호화폐 분석을 위한 인공지능 기반의 통합모델 제안, 한양대학교 공학대학원, 석사학위논문
- 천도현·김지훈·김병천(2017), “확률적 변동성 모형을 이용한 원화 환율 변동성 추정”, 통계연구(2017), 제22권 제3호, 68-99
- 한영희·서석환(2015). “상장기업의 재무적 특성과 기업가치 간의 상관관계에 관한 연구”, 2015년 (사)한국국제회계학회

추계국제학술발표대회, p460-476

황승환(2020), 전문기관의 원/달러 환율예측력 분석:시계열 모형 및 기계
학습 LSTM 모형과의 비교연구, 서강대학교 경제대학원, 석
사학위논문

<국외문헌>

- Basu, Sanjoy(1977). "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis." The Journal of Finance 32 (3). Wiley Online Library: 663 - 82.
- Boies, K. & Howell, J. M.(2006). Leader - member exchange in teams: An examination of the interaction between relationship differentiation and mean LMX in explaining team-level outcomes. The Leadership Quarterly, 17(3), pp. 246-257.
- Baker, Malcolm, Brendan Bradley, and Ryan Taliaferro(2014). "The Low-Risk Anomaly: A Decomposition into Micro and Macro Effects." Financial Analysts Journal 70 (2). Taylor & Francis: 43 - 58.
- Brunnermeier, Markus K, and Jonathan A Parker(2005). "Optimal Expectations." American Economic Review 95 (4): 1092 - 1118.
- De Bondt, Werner FM, and Richard Thaler(1985). "Does the Stock Market Overreact?" The Journal of Finance 40 (3).

- Wiley Online Library: 793 - 805.
- Engle, R., & Granger, C(1987). Co-integration and Error Correction:
Representation, Estimation and Testing.
Econometrica, 55(2), 251 - 76.
- Hsu, Jason, Vitali Kalesnik, and Engin Kose(2019). “What Is Quality?”
Financial Analysts Journal 75 (2). Taylor &
Francis: 44 - 61.
- Hyndman, R.J. & Athanasopoulos, G(2018) Forecasting: principles and
practice, 2nd edition, OTexts: Melbourne,
Australia. OTexts.com/f pp 2.
- Hyung jun Park, Min KyuSim Dong, GuChoi(2019),An intelligent
nancial portfolio trading strategy using deep
Q-learning , Expert Systems with Applications
Volume 158, 15 November 2020, 113573
- Brown, R. G.(1959). Statistical forecasting for inventory control.
McGraw/Hill.
- Holt, C. E(1957). Forecasting seasonals and trends by exponentially
weighted averages (O.N.R. Memorandum No. 52).
Carnegie Institute of Technology, Pittsburgh USA.
<https://doi.org/10.1016/j.ijforecast.2003.09.015>
- Jegadeesh, Narasimhan(1990). “Evidence of Predictable Behavior of
Security Returns.” The Journal of Finance 45 (3).
Wiley Online Library: 881 - 98.
- Jegadeesh, Narasimhan, and Sheridan Titman(1993). “Returns to

- Buying Winners and Selling Losers: Implications for Stock Market Efficiency.” The Journal of Finance 48 (1). Wiley Online Library: 65 - 91.
- Lehmann, Bruce N(1990). “Fads, Martingales, and Market Efficiency.” The Quarterly Journal of Economics 105 (1). MIT Press: 1 - 28.
- Rendleman Jr, Richard J, Charles P Jones, and Henry A Latane(1982). “Empirical Anomalies Based on Unexpected Earnings and the Importance of Risk Adjustments.” Journal of Financial Economics 10 (3). Elsevier: 269 - 87.
- Piotroski, Joseph D, and others(2000). “Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers.” Journal of
- Qian, Edward(2011). “Risk Parity and Diversification.” The Journal of Investing 20 (1). Institutional Investor Journals Umbrella: 119 - 27.
- Accounting Research 38: 1 - 52.
- Sharpe, William F(1964). “Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk.” The Journal of Finance 19 (3). Wiley Online Library: 425 - 42.
- Sharpe, William F(1964). “Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk.” The Journal of Finance 19 (3). Wiley Online Library: 425 - 42.

Sefton, James, David Jessop, Giuliano De Rossi, Claire Jones, and
Heran Zhang(2011). “Low-Risk Investing.” UBS
Investment Research.

Winters, P. R(1960). Forecasting sales by exponentially weighted
moving averages. Management Science, 6, 324 - 342.



ABSTRACT

A Study on DQN Reinforcement Learning Using Monte Carlo Simulation in Portfolio Configuration Using Quant Analysis

Eui-Cheol Kim

Department of U-City, The Graduate School

Seoul Venture University

Seoul, Korea

(Supervised by professor Chan-Gyu Hwang)

The basis of the era of the 4th industrial revolution is the development of intelligent (smart) algorithms along with the development of computer performance and networks. The word SMART, which represents this phenomenon, is an intelligence that immediately supplements or replaces the human role according to the situation through an automated algorithm for tasks that require human judgment and discernment beyond the wired-wireless connection of the existing computer computing power. (Intelligence) type system is installed. The basis of SMART is the development of statistical analysis and mathematical techniques. And these methodologies are developing in the direction of analyzing and utilizing

continuous time series data along with computer communication environments such as big data, IOT, and AI.

This study outlines representative time series analysis methods used in financial analysis and investment fields in this direction of development, and deep learning.

, And machine learning techniques such as reinforcement learning were applied to an actual stock investment portfolio analysis. Furthermore, it was examined to what extent the reinforcement learning model trained by generating virtual data, not real data, can learn about stock investment portfolios.

First, for this purpose, we first look at the capital asset pricing model (CAPM), the generalized autocorrelation conditional heteroscedasticity (GARCH) model, and the cyclic neural network (RNN) as the theoretical background for stock portfolio investment and time series analysis, and finally, DQN reinforcement learning and GBM We looked at Monte Carlo Simulation.

In order to implement this methodology as a model, information such as financial statements and stock prices of stocks were collected through a web crawling method. After that, they were organized into value indicators, financial indicators, and momentum indicators and used to select the components of the portfolio.

Second, the selected stock price was used to calculate the portfolio composition ratio using the time series analysis method. At this time, the portfolio composition strategy was to set the minimum volatility ratio strategy as the same strategic condition, and to examine the analytical characteristics of the four methodologies in actual stock price data.

Among them, DQN reinforcement learning creates two virtual data using

GBM Monte Carlo Simulation as well as real data, and learns them separately to examine the possibility of learning through virtual data.

Third, a back test was performed on the stock price data of three stocks showing different patterns for the models prepared with four methodologies for the comparison test. Portfolios were constructed from models composed in different ways, and the cumulative return results were compared and analyzed.

Through this comparison, we tried to confirm that the development direction of the intelligent algorithm used in the analysis method has been developed continuously in the model that calculates the recorded data value by standardizing it. First, it was developed as a method of setting an environmental model and predicting results from input variables, and it was found that it is developing as a machine learning method that finds relationships and features from the data itself.

Fourth, if the development of the SMART methodology introduces an algorithm that reflects the environment beyond the limits of the given data, we tried to examine the possibility of the SMART algorithm itself learning the predicted or possible change situation.

The results of this study, applied and compared with this methodology, are as follows. First, as a result of overviewing the methods of time series analysis and applying them to financial analysis, it was found that the analysis method has developed in the direction of proactively predicting and reflecting environmental factors. Next, the models generated from each method were applied to actual stock price data to compose a portfolio and compared the cumulative returns. As a result, the more advanced the analysis method was, the more predictively the model responded to the changes in the data. Finally, there are cases in which loss occurs as a result of applying the

model that learned the policy on the judgment of the stock price data and the portfolio composition ratio by using reinforcement learning 100 times to the actual stock price data, but in most cases, it was found that the cumulative rate of return generates profit. Could

Furthermore, even in models trained using virtual data, the number of cases in which profits were generated like models trained using only real data increased. However, it can be seen that the longer the length of the virtual data, the higher the rate of return is generated, while the more loss is generated, resulting in lower stability. This problem seems to be able to be stabilized through additional training, and it is expected to be more efficient than the time and effort to collect actual data.

The important result is that the random change in stock price is generated by an algorithm, and it is confirmed that the SMART algorithm can be learned through this. This means that the factors of change in the investment environment are not analyzed as historical data, but as an algorithm for the possibility, and policies for this can be learned with the SMART system.