**목차**

1. **프로젝트 개요**

1.1 프로젝트 주제 선정의 배경

1.2 프로젝트 주제 필요성

1.3 프로젝트 특징

1.4 프로젝트 범위

1.5 프로젝트 참여 인력

**2. 단계별 추진 결과**

2.1 수행 일정 및 주요 TASK

2.2 시스템 설계 단계

**3. 프로젝트 수행 성과**

3.1 프로젝트 구성도

3.2 데이터베이스 구성

3.3 주가 데이터 수집 및 가공

3.4 종목 추천 전략 구성

3.5 백테스트

3.6 웹페이지 구현

**4. 추가사항**

4.1 개선 사항

4.2 한계점

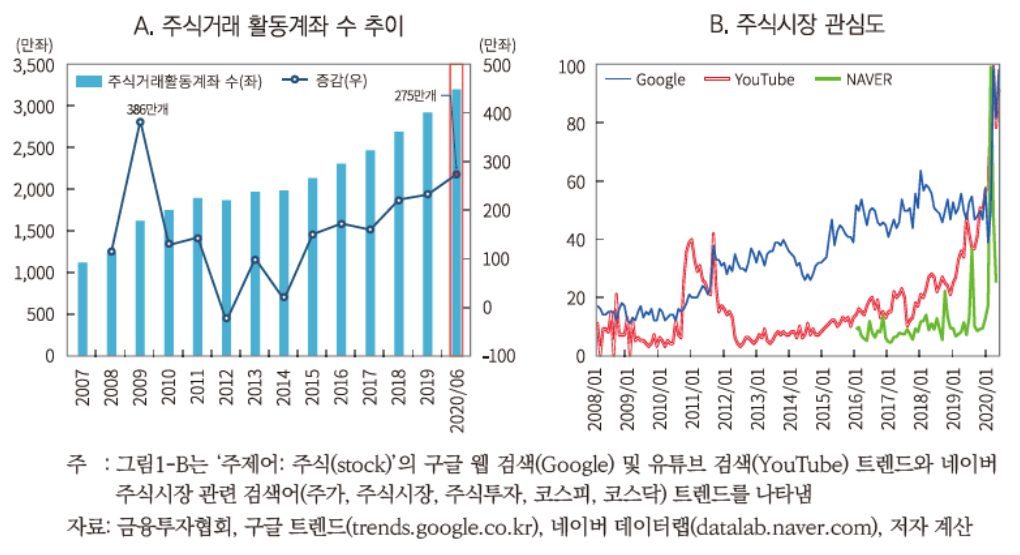
4.3 참고문헌

**1.프로젝트 개요**

**1.1 프로젝트 주제 선정의 배경**

최근 주식 거래 활동 계좌수가 꾸준히 증가중이며, 특히 작년 코로나 19사태 이후로, 신규 투자자의 수가 급증하고 있으며, 주식시장에 대한 관심도가 높아지고 있다. 남녀노소를 불문하고 투자에 대한 관심도가 상당히 증가하고 있는 추세이다. 이에 따라, 개인 투자자들을 위한 서비스를 생각하였다. 신규 투자자가 대폭 늘어나고 있는 상황이지만, 투자에 대한 경험이 없는 초보 개인투자자들, 소위 주린이라고 일컫는 투자자들을 위한 서비스가 많이 부족한 상황이다. 개인 투자자의 경우에는 충분한 정보를 얻지 못해 인터넷카페의 소문이나 종목토론방 등에 의존하여 종목에 대한 정보를 전혀 모른 채 매수를 하고 있는 경우가 비일비재하다. 어떤 종목이 좋은 종목인지, 좋지 못한 종목인지, 또한 투자를 하는데에 있어서 중요하게 생각되어야 할 요소가 무엇인지에 대해 잘 알지 못한 채 종목을 매수하는 경우가 많다. 이렇게 깜깜이 투자를 한 개인투자자는 손실을 보기 십상이다. 이러한 이유로, 개인투자자들에게 도움을 주고자 다음과 같은 주제를 진행하게 되었다. 앞으로 주식시장에 대한 관심은 꾸준히 지속적으로 증가할 것이라고 생각되며, 또한 종목의 재무데이터는 꾸준히 업데이트 될 것이다. 재무데이터가 꾸준히 늘어남에 따라 테스트 할 수 있는 데이터는 늘어나게 될 것이며, 장기적으로 보유하면 좋은 재무 상태가 좋은 우량주를 추천해주며 , 해당 종목들에 대한 정보를 제공해주고자 이와 같은 주제를

[그림 1] 꾸준히 오르는 주식시장의 관심도

선정하게되었다.

**1.2 프로젝트 주제 필요성**

본 설계에서의 궁극적인 목표는 개인 투자자들의 종목추천과 대시보드 제공을 활용한 정보제공이다. 주식 투자는 보통 기본적 분석과 기술적 분석으로 나눌 수 있다.

기본적 분석(Fundamental Analysis)의 경우에는 사업보고서, 재무제표, 기업의 Valuation을 분석하여 기업을 분석하며 투자하는 것이다.

기술적 분석(Technical Analysis)의 경우는 거래량, 지지선과 저항선, 추세 파악, 이동평균선, 보조지표 등을 활용한 투자이다. 기술적 분석의 경우는 보통 단기적 투자 관점에서 진행되며, 개인의 투자성향에 따라 많은 차이가 있으며, 꾸준한 수익률을 내는 데에 있어서 어려움이 많으며, 기업의 fundamental을 배제하여 투자를 진행하게 되는 경우가 많다.

본 설계는 기본적 분석(Fundamental Analysis)를 통한 분석으로, 기업의 재무데이터를 활용한다. 재무상태가 좋은 종목을 선정해줌으로써 종목 선정에 어려움을 겪는 개인 투자자들의 필요를 충족시켜주는 것을 궁극적인 목표로 진행한다.

**1.3 설계 특징**

본 설계는 수치만을 활용한다는 특징을 갖는다. 수치에 기반한 투자는 보통 Quant라는 단어로 많이 일컫게 되며, Quant란, Quantative Analysis의 약자로 계량적 분석을 활용한 투자를 의미한다. 본 설계는 재무데이터를 기반으로 한 기본적 분석(Fundamental Analysis)으로 진행한다. 재무데이터의 수집을 중점적으로 진행한다. 재무데이터의 경우 상당히 많은 정보들을 담고 있다. 가장 대표적으로 재무에 대한 정보를 제공해주는 사이트인 FnGuide 사이트에서 한국에서 가장 대표적인 주식인 삼성전자의 경우에만 하더라도 가치지표를 제외한 재무데이터의 경우 237개의 데이터가 존재한다. 해당 데이터를 모두 고려하여 머신러닝 기법을 활용하여 기업에 대한 스코어링 혹은, 1년후의 주가를 라벨로 하여 머신러닝 기법을 사용할 수 있다. 하지만, 종목별로 재무 데이터의 특성이 다르며 모든 데이터를 갖고있지 않다는 특징이 존재하였다. 이러한 이유로, Rule-based 기반의 전략을 구성하여, 구성한 전략을 활용하여 연도별로 종목추천으로 진행하였다. 또한 구성한 전략들은 **(참고문헌 참고)** 구성하였다. 또한 구성된 전략은 퀀트 투자에 있어서 유명한 reference들이 다수 존재하며, 기본적인 전략으로 구성하였다.

또한, 설계 초기에는 보조지표를 활용한 기술적 분석을 진행하였다. 하지만, 초기 설계서와

같은 전략을 통해 백테스트를 진행해본 결과, 그 결과가 생각보다 좋지 않아 기술적 분석을 배제하고 재무제표만을 활용한 기본적 분석(Fundamental Analysis)만을 진행한다. 기본적 분석을 통해, 재무상태가 좋은 우량한 종목을 선정해주며, 비교적 우량주가 많은 KOSPI 종목을 활용하여 진행한다.

**1.4 프로젝트 범위**

- 기반데이터 수집

- 보조지표, 재무데이터를 이용하여 데이터 가공

- 조건식 설정 및 알고리즘을 통하여 가공된 기반 데이터를 추가 가공 및 추천식 설정

- 웹을 통해 추천된 데이터들을 사용자에게 보여질 수 있도록 전체 종목 및 종목 상세 데이터 및 증권 정보, 시각화 자료 연계

- 웹 구현 및 추천식 설정을 위해 서버 및 DB 설정

**1.5 프로젝트 참여 인력**

분석 및 설계 : 김동호, 박병현

백엔드 및 서버 구축 : 김동호

알고리즘 설계, 데이터 분석 및 가공 : 박병현

1. **단계별 추진 결과**

**2.1 수행 일정 및 주요 TASK**

총 2개월 동안 프로젝트 설계 및 분석 수행후 개발 진행

* + - * 1. WBS

|  |  |  |
| --- | --- | --- |
| **Task** | **Owner** |  |
| **✔ 4/23 ~ 4/29** | | |
| AWS 서버 구축 | 김동호 |  |
| DB 구축 | 김동호 |  |
| 주식데이터 및 재무데이터 크롤러 구축 | 박병현 |  |
| 대시보드에 활용할 증권사 보고서 크롤러 구축 | 박병현 |  |
| **✔ 4/30 ~ 5/6** | | |
| 데이터 수집 자동화 및 DB연동???? | 김동호 |  |
| UI 설계 | 김동호 |  |
| 보조지표를 구하기 위한 모듈 생성 ex)RSI, MACD ...etc | 박병현 |  |
| DBSCAN 군집화를 활용한 지지선 생성 | 박병현 |  |
| **✔ 5/7 ~ 5/13** | | |
| UI 구현 | 김동호 |  |
| 대시보드 차트 시각화 | 박병현 |  |
| LSTM 종가예측 모델 생성 | 박병현 |  |
| **✔ 5/14 ~ 5/20** | | |
| 대시보드 구현 및 시각화 연동 | 김동호 |  |
| 대시보드에 활용할 관련주식 뉴스 RSS 활용하여 연동 | 김동호 |  |
| 보조지표를 활용한 종목추천 조건식 함수 생성1 | 박병현 |  |
| **✔ 5/21 ~ 5/27** | | |
| 조건식 함수와 웹에서의 버튼 연동하기 | 김동호 |  |
| UI를 통한 테스트 | 김동호 |  |
| 보조지표를 활용한 종목추천 조건식 함수 생성2 | 박병현 |  |
| **✔ 5/27 ~ 6/3** | | |
| UI 개선 및 성능 개선 | 김동호 |  |
| 대시보드 기능 개선 | 박병현 |  |
| **✔ 6/4 ~ 6/10** | | |
| UI 개선 및 성능 개선 | 김동호 |  |
| 추천된 종목 백테스트 구현 | 박병현 |  |
| **✔ 6/11 ~ 6/17** | | |
| 마무리 |  |  |
| 마무리 |  |  |

* + - * 1. 주요 마일스톤



**2.2 시스템 설계 단계**

- 작업환경

① 개발 환경

✔ 데이터 수집 도구 : R(R Studio), Python(Pycharm) , BeautifulSoup

✔ 데이터 정제 및 EDA : jupyter notebook, pandas, numpy, matplotlib

✔ 시각화 : matplotlib, pyplot , plotly, yahoo\_finance(python library)

✔ 활용 서버 : AWS EC2

✔ 활용 DB : Maria DB(AWS RDS)

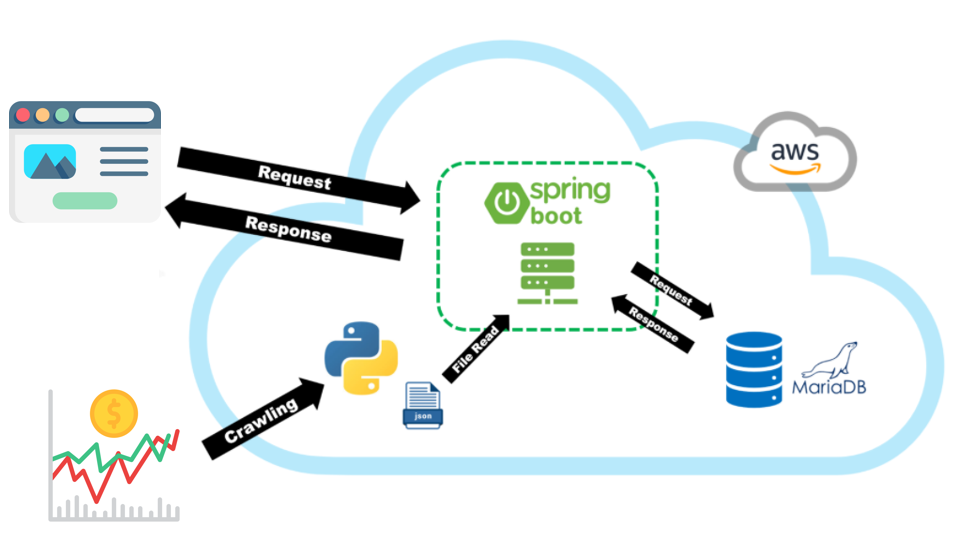
✔ 배포 및 자동화 :Travis CI

✔ 웹개발 : Spring Boot (JAVA 11)

✔ BUILD : Gradle (4.10.2)

② 운영 환경

Aws ec2 를 이용하여 웹서비스 구현 및 travis ci 배포 자동화



[그림 2] 시스템 설계 구조 요약

**3. 프로젝트 수행 성과**

**3.1 프로젝트 구성도**

본 프로젝트의 진행사항은 크게 6가지로 나눌 수 있다.

- 데이터 수집

- 종목 추천 전략 구성

- 전략을 활용한 종목 추천

- 전략 별 백테스트

- 시각화 및 대시보드 구현

- 스프링부트를 통한 개발 및 AWS를 이용한한 웹 서비스 구현

첫 진행과정으로는 주식 정보 및 주가 데이터 수집을 진행한다. 수집하는 주식의 정보로는 기업의 가치지표, 재무제표, 기업 설명을 수집한다. 그 후, 기업별로 수집한 데이터를 쉽게 활용하기 위하여 데이터 정제 과정을 진행한다. 기업 설명은 고정적인 정보이며, 가치지표와 재무제표같은 경우에는 주기적으로 업데이트 되지만, 그 주기가 3개월이다. 본 설계에서는 연도별 재무데이터만을 활용하기 때문에, 업데이트 주기가 1년이므로, 재무데이터 수집의 경우는 자동화를 진행하지 않는다.

그 후, 수집 후 가공한 재무데이터를 토대로 종목 추천 전략을 구성하였다. 종목 추천 전략은 모두 재무데이터와 가치지표만을 활용한다. 이 때 사용한 전략들은 (참고문헌 참고) 구성하였으며, 사용한 전략으로는 PER, ROE를 활용한 마법공식 전략, 가치지표를 활용한 Value 전략, 현금흐름과 유동성 및 기업의 성장성이 고려되는 F-Score전략을 활용하였다. 세 가지 전략을 활용하여, 연도별 12월 재무데이터를 활용하여 10가지 종목 추천을 진행한다. 해당 연도 12월 재무데이터를 활용하여 추천된 종목 10개를 모두 매수하엿을 경우, 그 다음 해 1년간 소지하고, 말일에 매도하였을 경우에 대한 백테스트를 진행하였다. 백테스트의 케이스는 3가지 전략으로 2017, 2018, 2019, 2020 각 연도의 12월 재무데이터를 활용한 그 다음해의 수익률을 통한 백테스트로, 총 12가지의 케이스밖에 없다는 한계를 갖고 있다. 초기 각 분기별로 진행을 할 예정이였지만, 재무데이터를 활용한 투자를 하는 경우는 가치투자로써, 최소 1년은 보유하여야 그 성과가 나타나는 경향이 있다. 그러므로 1년의 연도별 데이터를 활용하였다. 백테스트를 통해 해당 전략들을 매수하였을 경우에 대한 수익률을 python 라이브러리인 Plotly를 통하여 코스피 지수 대비 수익률을 확인할 수 있도록 구현하였다. 각 연도별, 전략별 추천 종목과 해당 종목을 추천받아, 해당 종목으로만 포트폴리오를 구성하였을 경우, 그때의 수익률을 웹을 통하여 확인할 수 있도록 웹을 구현하였다.

웹서비스는 aws 서버를 통하여 서버 구축 및 운영을 하며, 스프링부트를 이용하여 개발을 진행한다. 추천식이나 가공 데이터와 시각화 자료들을 웹을 통해 사용자에게 보여질 수 있도록 구성하며 사용하고자 하는 전략에 대한 전반적인 종목들과 그 종목에 해당되는 시각화 자료나 종목 상세 대시보드를 통해 사용자가 접근 할 수 있도록 한다.

**3.2 데이터베이스 구성**

- 테이블리스트

|  |  |  |  |
| --- | --- | --- | --- |
| 구분 | 테이블 ID | 테이블 명 | 비고 |
|  | BASE\_DATA | BASE\_DATA | 크롤링한 매일매일의 기반데이터(종가) 들을 관리하기 위한 테이블 |
|  | FSCORE\_DATA | FSCORE\_DATA | 전략별로 추천된 종목들을 관리하기 위한 테이블 |
|  | MAGIC\_DATA | MAGIC\_DATA | 전략별로 추천된 종목들을 관리하기 위한 테이블 |
|  | VALUE\_DATA | VALUE\_DATA | 전략별로 추천된 종목들을 관리하기 위한 테이블 |

- 칼럼구성도

|  |  |  |  |
| --- | --- | --- | --- |
| 구분 | 테이블 ID | 테이블 명 | 칼럼 구성 |
|  | BASE\_DATA | BASE\_DATA | Id,code,date,title,target,opinion,writer,source,link |
|  | FSCORE\_DATA | FSCORE\_DATA | Id,code,roa,cfo,accurual,lev,margin,turn,name,sector,price,pbr,pcr,per,psr,definition |
|  | MAGIC\_DATA | MAGIC\_DATA | Id,code,per,pbr,psr,pcr,roa,name,rank,sector,price,description |
|  | VALUE\_DATA | VALUE\_DATA | Id,code,pbr,psr,per,pcr,rank,name,sector,price,per,description |

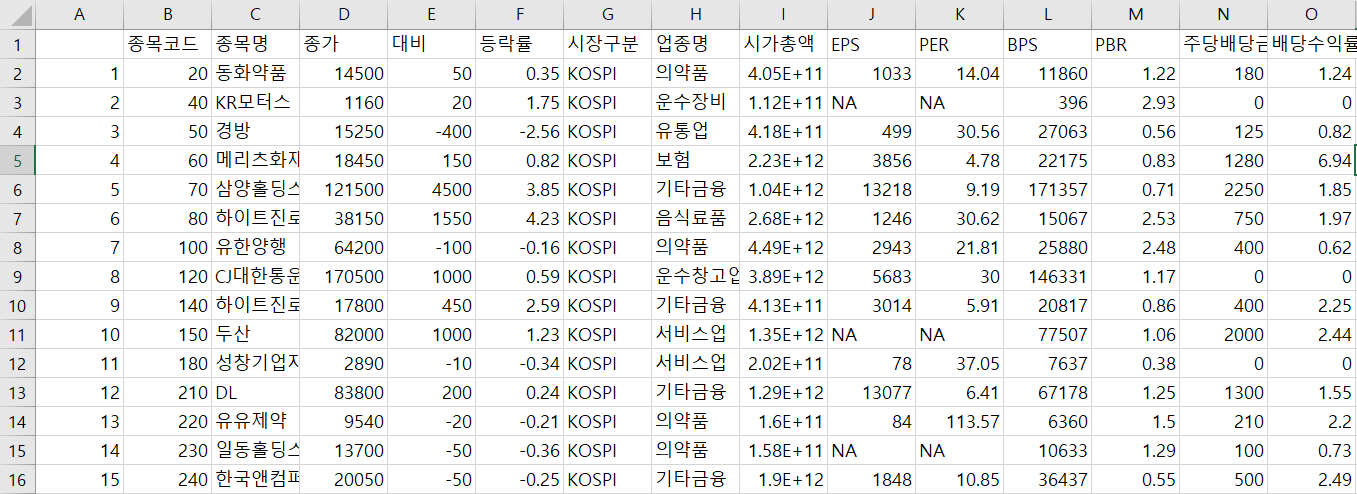
**3.3 주식 데이터 수집**

**✔ KOSPI 종목 783개의 종목 활용**

KOSPI 상장 종목은 약 900개의 종목이 존재한다. 하지만, 우선주와 스팩주를 제외하였다. 우선주의 경우에는 대신 보통주를 활용하는 것으로 해결 가능하며, 스팩주의 경우에는 유동성이 크다는 특징을 갖고 있으므로, 재무데이터를 활용한추천이 부적절하다고 판단하였다. 또한, 본 프로젝트는 재무상태가 우량한 종목을 추천하는 것을 목표로 하므로 코스닥 종목보다 코스피 종목을 활용하였다.

**✔ KOSPI ticker 수집**

재무데이터는 783개의 KOSPI종목의 재무데이터를 KRX 정보데이터 시스템 사이트(http://data.krx.co.kr)에서 수집하였다. R을 활용하여 수집을 진행하였다. 종목코드와 업종, 배당금에 대한 정보를 담고 있는 KOSPI\_ticker.csv 파일을 생성했다.

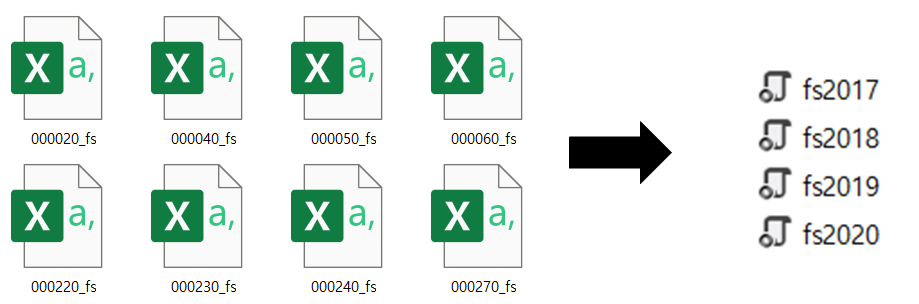
　＊Script : [~/R\_scripts/1. KOSPI\_ticker.R]

**[그림 2] KOSPI\_ticker.csv 예시**

**✔ 재무데이터 수집**

재무데이터는 분기별 데이터가 아닌 연간 재무데이터를 수집한다. 데이터 수집은 FnGuide사이트로부터, R을 활용하여 수집을 진행한다. 783개의 종목을 CSV로 저장 =>　python(pandas)을 활용한 연도별로 정제한다.

\*Script : [“~/R\_scripts/2. KOSPI\_ticker.R“, “~/KOSPI\_ticker.ipynb“,

“~/1. 연도별 재무데이터프레임 생성.ipynb”]

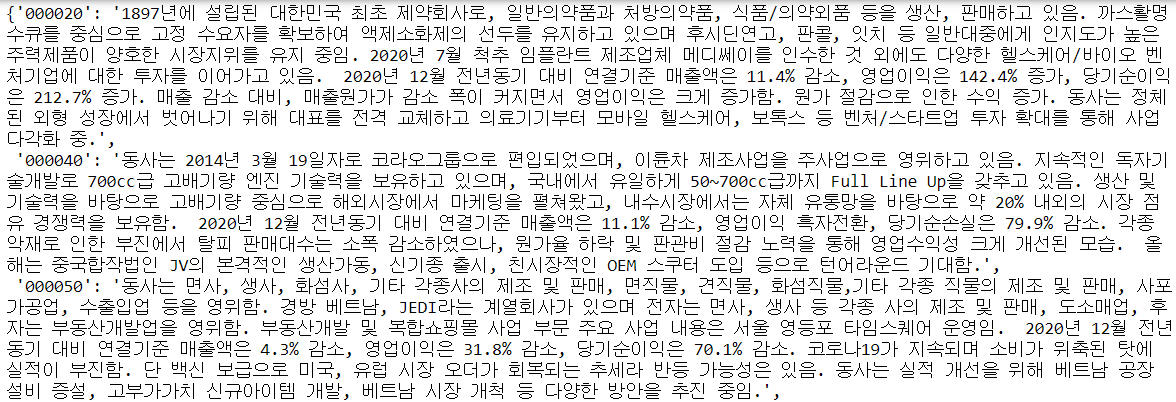
**개별 종목의 재무데이터.csv => 연도별 재무데이터.json**

**✔ 종목의 기업 설명정보 크롤링**

대시보드에 활용할 기업정보 크롤링 후, stock-description.json파일로 기업정보 저장.

python을 활용하여 Fnguide사이트 Requests, BeautifulSoup4 활용한 크롤링

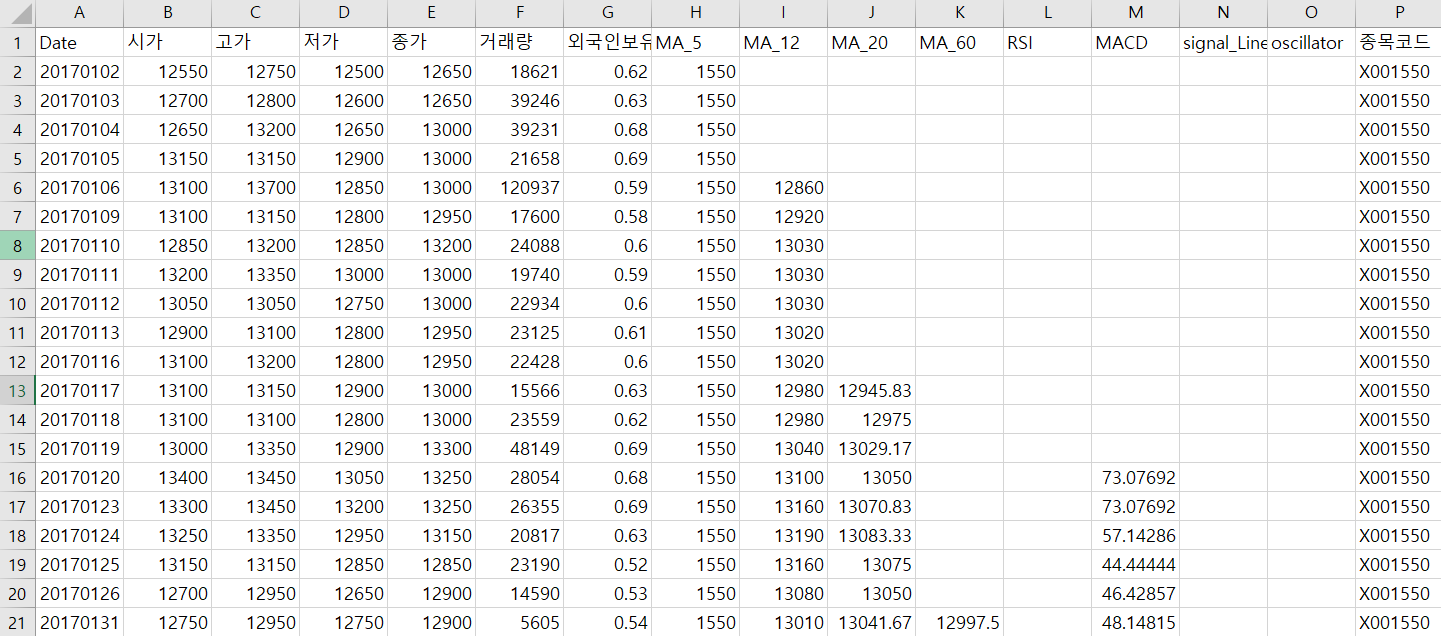
\*Script :　[~/기업정보 크롤링.ipynb]

****저장 경로 : [~/stock-description.json]]

[그림 3] 기업정보 json파일 예시

**✔ 종목별 주가 수집**

종목별 대시보드에 활용될 일봉 차트를 위하여 종목별 주가 수집을 진행한다. python을 활용하여 주가데이터 수집을 진행한다. 주가 데이터 수집의 경우는 네이버 금융 사이트 (<https://finance.naver.com)에서> 수집을 진행한다. 783개의 KOSPI 종목에 대해서 수집을 진행하며, 수집하는 항목은 일봉 차트를 그리는데에 활용하기 위한 [시가, 저가, 고가, 종가, 거래량] 등의 데이터를 수집한다. 데이터 수집은 2017년~ 현재데이터만을 수집한다. 또한, 첫 설계에서는 기술적 분석을 진행 예정이였으므로, 해당 기술적 분석을 진행 활용을 하기 위한 보조지표 (RSI, 이동평균선, MACD) 함수를 구성하여, 주가를 수집과 동시에 해당 보조지표값을 계산한다.

\*Script : “~/python-stock-crawler.py”

[그림 4] 주가 정보 수집 데이터

**3.4 종목 추천 전략 구성**

종목 추천 전략은 크게 3가지로 구성된다. 모든 종복 추천 전략은 각 연도별 12월 재무데이터를 기반으로 종목 추천을 진행하며, 추천 종목 수는 설정할 수 있도록 구현하여 언제든 변경 가능하지만, 각 전략 모두 추천종목을 10가지 종목으로 한정하여 진행한다.

1. **마법공식(Magic Strategy)**

* 재무제표의 PER, ROA를 활용한 추천

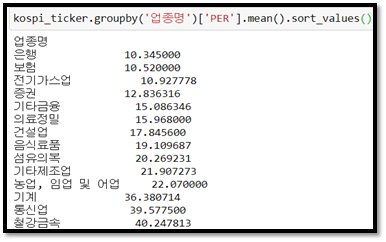
1. **F-Score 전략**

* 총 9가지의 재무상태 조건을 설정하여, 각 조건을 만족하면 +1점을 하여 0~9값을 갖으며, 만점의 값을 갖는 종목 추천

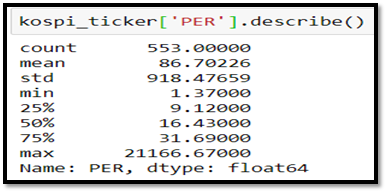
**3. Value 전략**

=> [PER, PSR, PBR, PCR] 전략을 활용한 추천

**3.4.1 마법공식 전략**

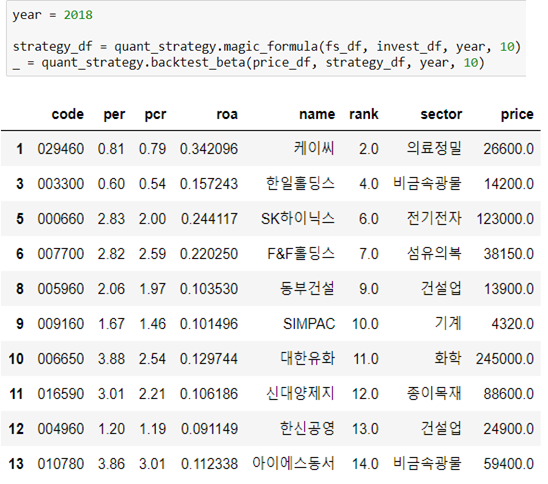
마법공식 전략은 PER, ROE값을 활용한다. PER이란, 현재주가를 EPS(주당 순이익)으로 나눈 값으로, 즉 시가총액을 당기순이익으로 나눈 값이다. PER이 낮을수록 저평가된 종목임을 알 수 있다. ROE는 당기순이익을 자본총액으로 나눈 값으로, 값이 높을수록 기업의 순 이익이 높다고 판단할 수 있다. PER은 대표적으로 자주 활용되는 기업의 가치를 판단하는 지표이다. 하지만, 종종 오차를 범하는 경우가 많으므로 ROE와 같이 활용하면 더 좋은 성과를 나타낼 수 있으므로, 이 두 지표를 같이 활용하는 전략이다. 해당 전략은 저 PER, 고ROA 종목을 뽑아내는 것으로, 783개의 종목에 대해서 PER, ROA 순위를 매긴다. PER은 오름차순 rank, ROA는 내림차순 rank로, 두 rank의 합산이 낮을수록 추천의 우선순위가 높아지는 rule-based기반의 전략이다.

[그림 5] 업종별 PER 평균값

해당 전략을 하기에 앞서, 종목별 PER을 살펴보니, 업종별로 차이가 크다는 점을 확인할 수 있었다. 은행, 보험, 증권, 전기가스업, 기타금융 업종의 경우 업종의 특성상 PER이 낮은 경향을 보인다. 또한, 금융과 관련된 주식들은 다른 업종들과 다른 특성을 보이는 경우가 많으므로, 해당 업종을 갖는 종목들에 대해서는 추천을 제외한다.

[그림 6] 783개 KOSPI 종목 PER 분포

또한, PER값에 대한 분포를 살펴보니, 상당히 큰 값을 갖는 경우도 존재하며, min값이 1.3으로 매우 낮은 수치를 보이는 경향도 존재하였다. 이 외에, 다른 보조지표들 또한 회계상의 이유로 이상치가 종종 존재하는 것을 확인할 수 있었다. 따라서, 추천을 하는데에 있어서는 사용되는 지표가 모든 업종에서의 1%이하 혹은 99%이상인 값을 갖는 경우에 대해서는 추천대상에서 제외하기

로 하였다. 이렇게 추천된 종목들을 확인해 볼 수 있다.

[그림 7] 마법공식 종목 추천 예시

[그림7]는 2018년12월 재무데이터로부터 추천된 종목의 예시이다. 연도를 2020으로 설정한다면 , 올해 추천받을 종목이 나온다. 다음과 같은 10가지의 종목이 추천되는 것을 확인할 수 있다. 각 종목을 모두 매수하였을 때 백테스트를 활용한 전략에 대한 검증또한 가능하다.

\* Script : [~/백테스트 구현.ipynb] , [ “~/quant\_startgey.py” => magic\_formula 함수]

**3.4.2 F-Score전략**

본 전략 또한 재무데이터를 활용한다. F-Score 전략은 총 9가지의 조건식을 활용하며, 각 조건식을 만족할때마다 1점씩 부여되어, 0~9점의 값을 갖게 된다. F-Score 전략은 기업의 수익성, 재무성과, 운영 효율성을 고려한다. 해당 전략은 지배주주순이익, 자산, 장기차입금, 유동자산, 유동부채, 매출액, 매출총이익, 매출액증가 등 다양한 요소가 활용된 전략이다. 또한 기본적 분석(Fundamental Analysis)에서 자주 활용되는 대표적인 기법이기도 하다.

✔ ***Fscore전략 9가지 조건식***

1. 당기 순이익이 0이상인가?

2. 영업활동현금흐름이 0 이상인가?

3. 영업활동현금흐름이 당기 순이익보다 큰가?

4. 부채비율이 전년도보다 감소하였는가?

5. 유동비율이 전년도보다 증가하였는가?

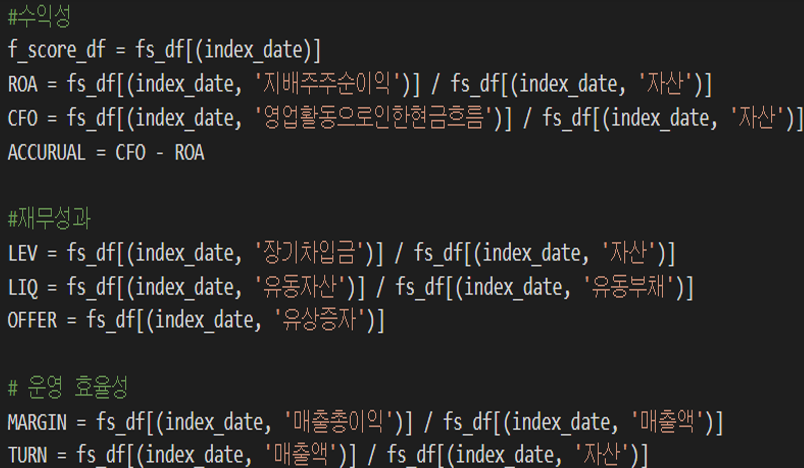
6. 총자산회전율이 전년도보다 증가하였는가?

7. 유동주식수는 전년도보다 증가하였는가?

8. 매출총이익은 전년도보다 증가하였는가?

9. ROA가 전년도보다 증가하였는가?

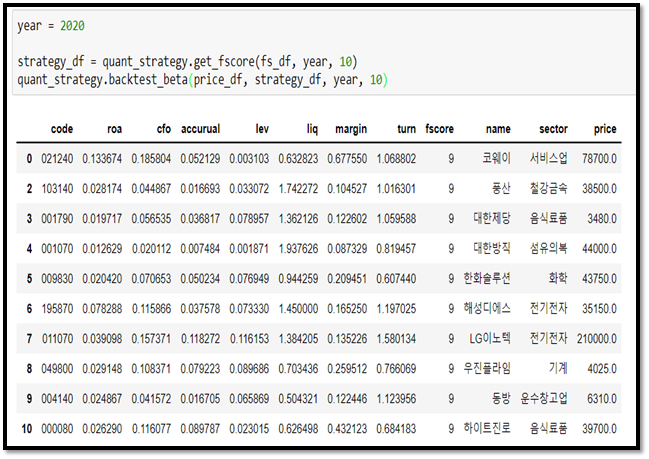
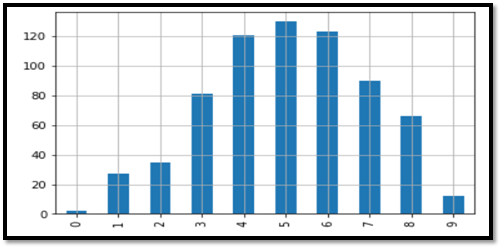
[참고문헌 : Value Investing: The Use of Historical : Joseph D. Piotroski]

각 조건을 만족할 때, 1점씩 부여된다. 1점씩 부여됨으로써, 0~9점의 값을 갖으며, 9점의 값을 갖는 종목은 재무상 좋은 종목으로 판단할 수 있다. 해당 전략또한 2017~2020 년도의 12월 재무데이터를 활용하여 추천을 진행한다.

파생변수 설명

다음과 같은 파생변수를 만들어, 생성한 파생변수를 활용하여 각 기업들이 9가지 조건식을

만족하는지 확인하며, 각 조건을 만족할 경우 +1점을 추가하여 종목별 F-Score값을 부여한다.

[그림 9] 는 2020-12년 재무데이터 기반으로 F-Score값이 9점을 갖는 종목이다. 코웨이, 풍산, 대한제당, 대한방직 등의 종목이 추천되었으며, 앞서 생성한 파생변수 기반으로 계산된 F-score값을 기반으로 추천이 되었다.

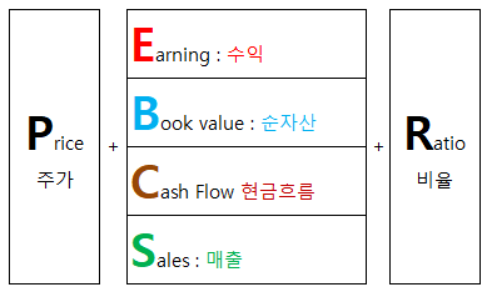
[그림 10] 783개 종목 F-score값 분포도

[그림 9] Fscore를 활용한 2020년 12월 재무데이터 기반의 추천 종목

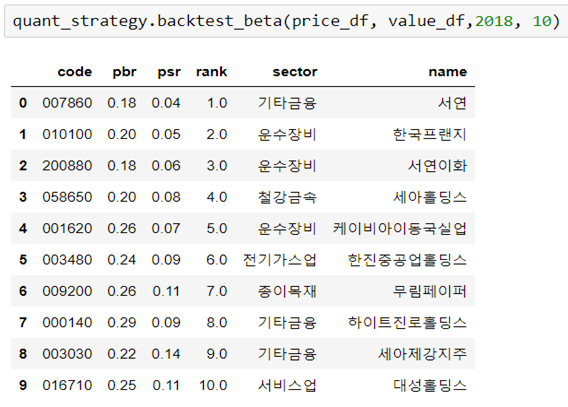
[그림10]은 2020-12년 재무데이터 기반의 F-Score 분포를 확인한 결과이다. F-Score값은 0~9점의 범위를 갖고 있으며 정규분포 그래프와 유사한 분포를 갖는다는 특징이 있어 대부분 평균값인 5점에 가까이 모여있는 것을 볼 수 있고, 0점에 가까울수록 재무상태가 좋지 않으며, 9점에 가까울수록 재무상태가 좋다고 판단할 수 있다. 본 전략에서의 추천은 9점인 종목을 대상으로 추천을 진행한다.

\* Script : [ “~/quant\_startgey.py” => get\_fscore 함수]

**3.4.3 Value 전략**

Value 전략은 가치지표를 활용한 전략이다. 기업의 가치지표로 활용되는 지표는 PER, PSR, PCR, PBR등의 지표가 존재한다.

가치지표 설명

수집한 PER, PBR, PCR, PSR 값을 모두 고려한 종목추천을 진행한다. 하나의 가치지표만을 활용하는 것보다, 여러개의 지표를 결합하는 경우 기업의 재무 특성을 더 고려할 수 있기 때문에, 모든 지표를 활용하여 추천을 진행한다. 4가지 지표들 중 사용자가 특정 지표를 조합하여 종목추천을 진행가능하도록 모듈화를 하였지만, 구현에서의 종목추천은 4가지 지표의 순위를 모두 고려하여 추천을 진행한다. 해당 전략에서도 상위 1%, 하위99%의 값을 넘는 값들은 이상치로 고려하여 추천 대상에서 제외한다. 본 전략은 783개의 종목들에 대해서 각 지표들의 순위를 매긴다. PER값만 낮을수록 저평가인 기업임을 의미하므로, PER값은 오름차순으로 순위를 매기며, 나머지 변수들에 대해서는 내림차순으로 높은 값일수록 낮은 rank값을 부여한다. 4가지 지표의 rank의 값을 합산하여 낮은 rank순으로 10가지 종목에 대해서 추천을 진행한다.

[그림 10] Value전략을 활용한 10가지 종목 추천

\* Script : [ “~/quant\_startgey.py” => make\_value\_combo 함수]

**3.5 백테스트**

\*Script : “~/quant\_startgey.py” => backtest\_beta 함수

앞서 ~3.4 까지는 종목 추천을 진행하였다. 각 전략들의 추천에 대한 결과를 검증하기 위한 백테스트를 진행한다. 백테스트는 절대적인 수익률을 확인함과 동시에 코스피 대비 수익률을 비교할 수 있도록 구현하였다. KOSPI 종목인 만큼, KOSPI 지수보다 높은 수익률을 갖는 전략들은 비교적 성공적인 전략이라고 판단할 수 있기 때문이다. 여기서 KOSPI 지수는 크롤링을 진행하지 않고, yahoo\_finance 라이브러리를 활용하여 KOSPI 지수를 불러온다.

본 설계에서 설정할 수 있는 연도는 2017/2018/2019/2020 년도로, 각 년도의 12월 재무데이터를 기반으로 1년간 투자를 시나리오로 진행한다.

그러므로, 각 연도의 12월 재무데이터를 기반으로 추천된 10개의 종목을 다음연도 1월~ 12월 1년간 보유하였을 경우 코스피 대비 수익률을 백테스트로 진행한다. 종목에 대한 보유비중과, 현금보유비중 등을 고려하여 구현하고자 하였지만, 모두 같은 비율을 갖는 것으로 가정하여 백테스트를 진행한다. 백테스트는 수익 금액보다, 수익률을 테스트 하는것에 중점을 맞춰 진행하였으며, 그렇기 때문에 초기자금은 100만원으로 고정시켜서 진행하였다. 또한, python 시각화 라이브러리인 plotly라이브러리를 활용한다.

plotly 라이브러리는 동적인 사각화 자료를 html형태 혹은 json형태로 저장하여 다양한 용도로 사용될 수 있다는 장점을 갖고 있다. 특정 구간의 수익률을 동적으로 파악해 볼 수 있기 때문에 plotly라이브러릴 활용한 시각화를 웹을 통해 제공한다.

백테스트는 앞서 다뤘던 3가지 전략인 마법공식, F-score, Value전략의 2017~2020 년도에 대해서 진행한다.

\*Script : [~/const.py => plot\_to\_json 함수 , “~/결과json/”]

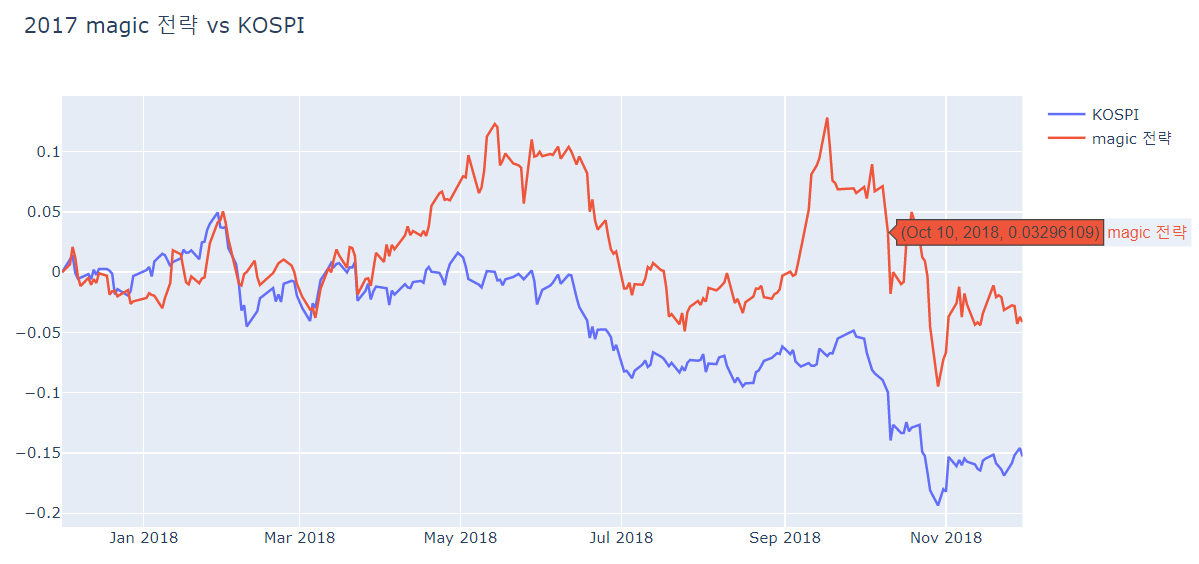


그림 11

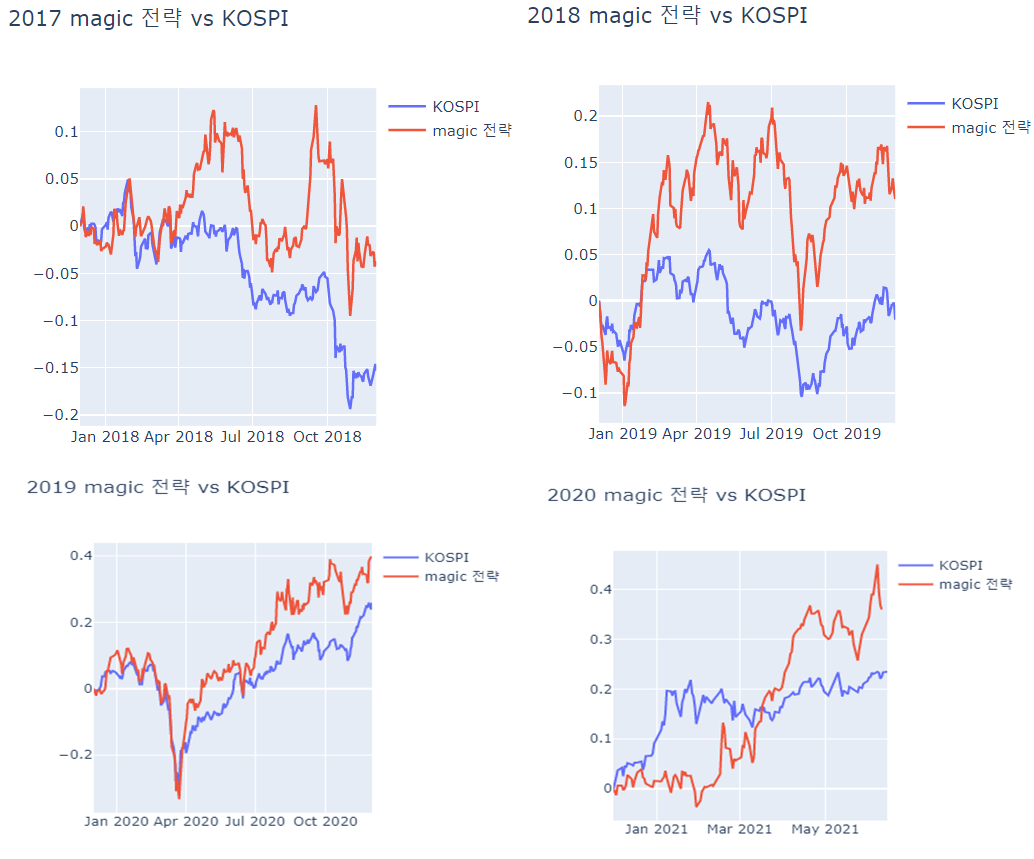
[그림 12] 마법공식을 통해 추천된 10가지 종목 수익률

위 [그림 12] 는 2017년도의 마법공식을 통해 추천된 10가지 종목을 매수하였을 경우, 코스피 대비 수익률에 대한 백테스트 결과를 plotly 라이브러리를 활용하여 시각화 한 그래프이다. 해당 그래프를 웹에 올리기 위하여 json파일로 저장하였고, 저장한 json은 [“~/결과json/“] 경로에 저장하여 사용한다. 다음과 같은 3가지 전략에 대한 추천종목의 결과를 확인할 수 있는 백테스트 시각화를 모두 json파일로 저장하여 사용한다.

**3.5.1 백테스트를 활용한 전략 검증**

본 설계에서의 한계는 연간 재무데이터가 2017~2020으로 4개 년도의 데이터만 활용한다는 한계가 존재한다. 만약 더 과거의 데이터를 구할 수 있다면 해당 전략들에 대한 신뢰성이 더 높아졌을 것이다. FnGuide를 수집한 데이터는 2016년 이전의 데이터가 결측치가 상당히 많이 존재하였기에, 2017년 이후의 데이터를 활용한다. 먼저 3개 전략의 수익률을 2017~2020년도, 4개년도를 시각적으로 비교해보도록 한다.

**3.5.2 마법공식 전략 백테스트**



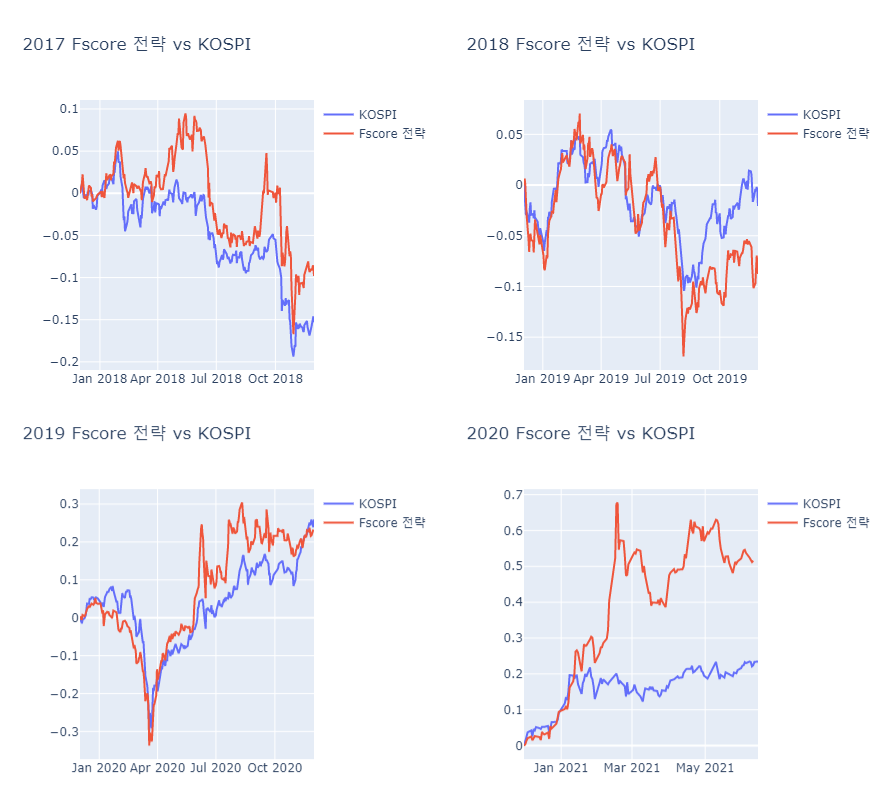
[그림 13] 마법공식 4개연도 수익률

2018년을 제외하고는 항상 수익률이 20%이상을 보이는 것을 확인할 수 있다.

**3.5.3 Value전략 수익률 백테스트**

[그림 14] Value 전략을 활용한 수익률

Value전략을 통해 추천된 종목 또한 2018년을 제외하고는 모두 코스피보다 수익률이 높으며, 수익률이 20%정도 나오는 것을 확인할 수 있다.

**3.5.4 F-score전략 백테스트 시각화**

[그림 15] F-score전략 4개년도 수익률

F-score전략의 경우에는 수익률이 비교적 저조한 것을 볼 수 있다. 각 전략들은 재무데이터를 활용한다는 한계를 갖는다. 공통적으로 2020-12 재무데이터를 활용하여 올해 추천되는 종목들의 수익률이 높은 것을 확인할 수 있다. 그 이유는 올해는 금리 인상의 이슈로 인해 재무상태가 좋은 가치주(우량주) 위주의 상승장이 진행되었기 때문이다. 그리고 재무상태가 좋더라도 현재 매수 가격이 비싸다면, 수익률을 보장하기 힘들다는 한계를 갖고 있다. 추천된 종목들을 보조지표를 활용하여 더 저렴한 가격에 매수하고, 적당한 시기에 매도를 하게 된다면 수익률은 더 극대화될 것이다. 세가지 전략 중에서는 마법공식전략이 가장 단순한 로직을 가지고 있지만, 비교적 수익률이 가장 좋은 것을 볼 수 있다.

**3.6 웹페이지 구현**

**1. 기반 데이터들은 밑과 같이 스케쥴러를 통하여 특정 시간에 데이터화 시킨다.**

AWS Linux 서버 내 Scheduling 설정도 진행한다.

@Scheduled(cron = "0 30 16 \* \* \*", zone = "Asia/Seoul")

public ResponseEntity<String> bulkUpdate()

throws JsonParseException, JsonMappingException, IOException {

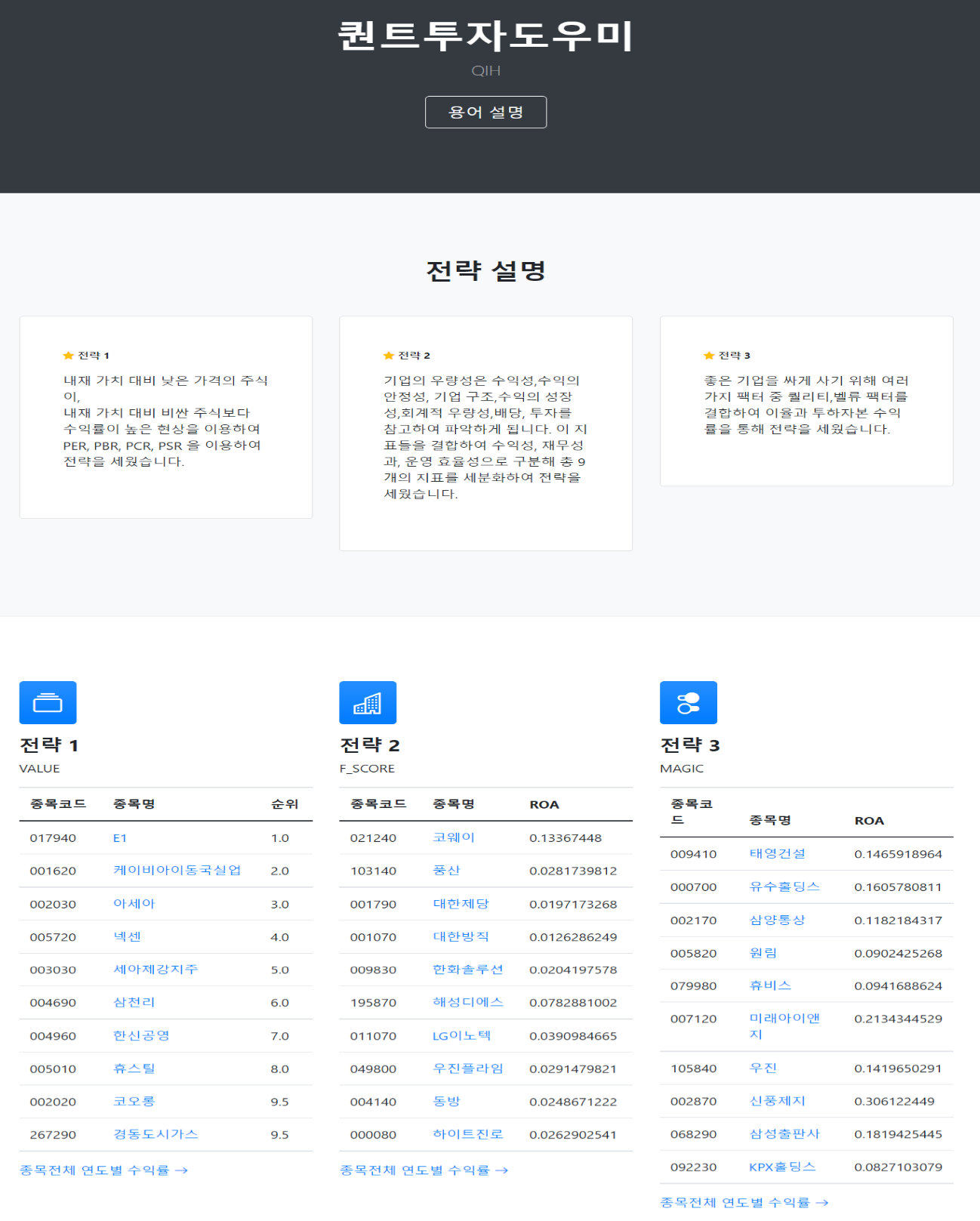
//해당 Logic...

}

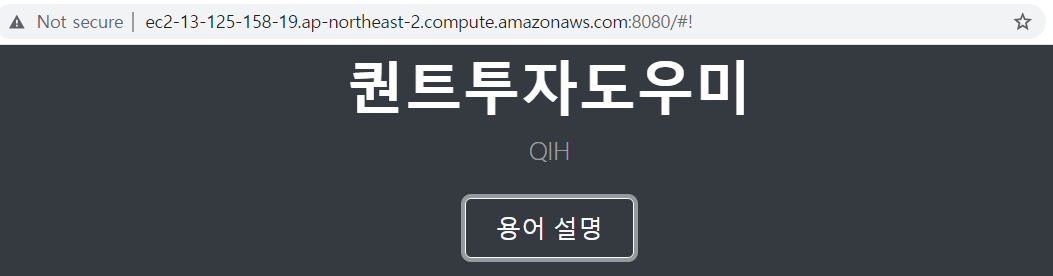
**2. 상세 구현 내용**

**실제구현 서비스 주소(**[**http://ec2-13-125-158-19.ap-northeast-2.compute.amazonaws.com:8080**](http://ec2-13-125-158-19.ap-northeast-2.compute.amazonaws.com:8080)**)**

css는 부트스트랩을 사용하여 화면 디자인을 구성하였고 사용자 화면(template)은 mustache를 이용하여 구성하였다.

**[****그림 16] 웹페이지의 메인 화면**

웹페지이믜 메인 화면은 [그림 16] 와 같다. 전략에 대한 이해를 돕기 위하여 전략에 사용한 지표들에 대한 용어들에 대한 정보를 제공해주고자 용어설명 팝업을 만들었으며, 메인에 전략 설명을 통해 전략에 대한 간단한 정보를 제공한다. 그리고 하단에 추천된 종목들은 각 전략으로 2020-12월의 재무데이터를 기반으로 추천된 종목으로, 올해 매수를 추천하는 종목으로 구성되어 있다.

**3. 용어설명 팝업**

앞서 전략에 사용한 재무데이터를 활용한 투자 지표들에 대한 용어에 대한 설명을 설명해주고자, “용어 설명“ 팝업을 구현. 해당 버튼을 누르게 되면, 용어설명에 대한 정보를 담고있는 팝업으로 이동하게 된다. PBR, PSR, PCR, PER, ROA, CFO 각 용어들에 대한 설명을 제공한다.

**4. 전략설명**

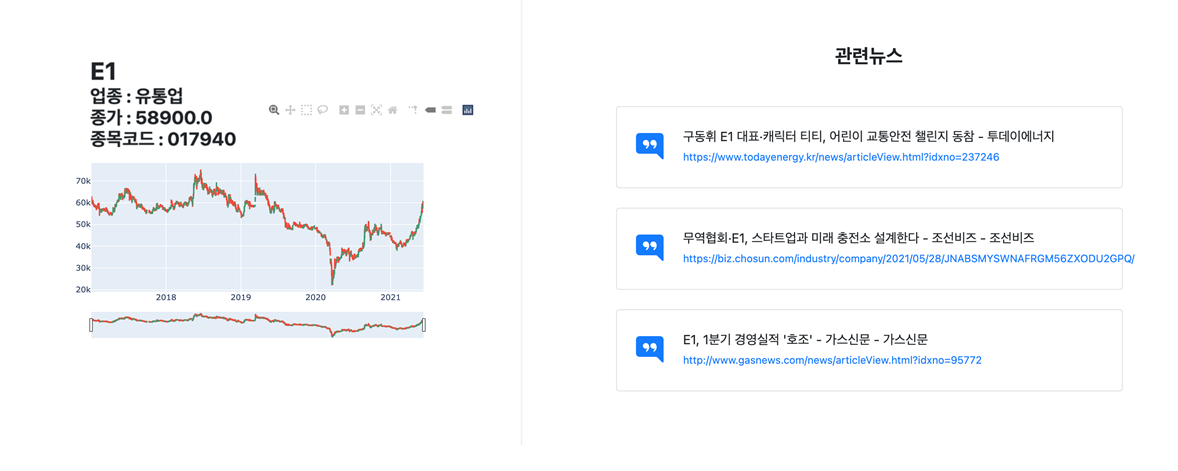
[그림16] 에서와 같이 웹 페이지에서 간단한 하게 각 전략에 대한 설명을 기술한 화면을 제공한다.

**5. 종목 추천 화면**

[그림16] 하단에서 볼 수 있듯이 3가지 전략에 대해서 각각 10가지 종목을 추천한다. 종목코드와 종목명과 같이 제공되며, 추천 종목의 개수는 전략별로 상위 10개로 추천한다.

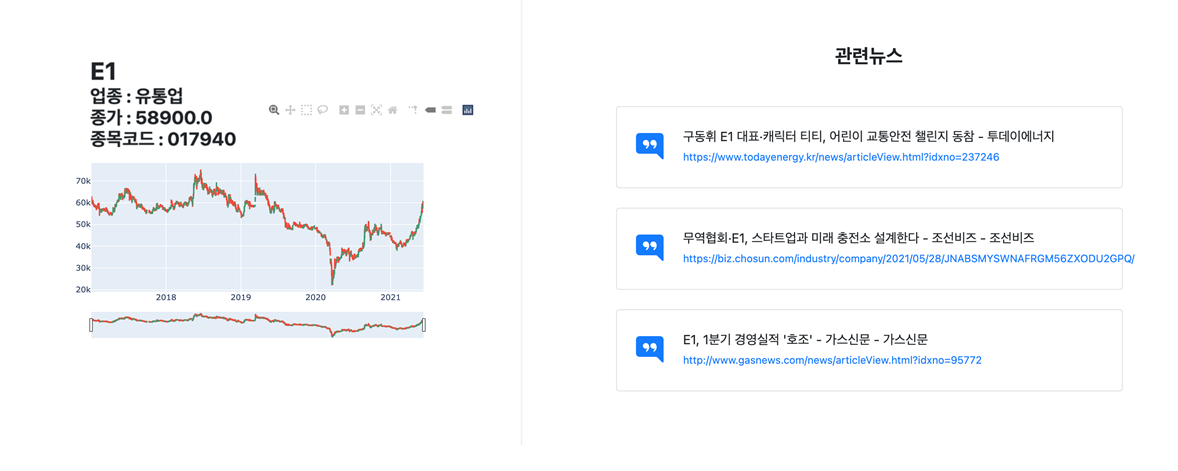
**6. 대시보드 화면**

위의 종목 추천 화면에서 추천된 종목명을 누르게 되면 종목에 대한 정보를 파악할 수 있는 대시보드로 이동하게 된다. 대시보드에서 제공하는 정보는 기업에 대한 정보를 제공한다.



**[그림 17] 기업의 대시보드 구성 1**

[그림 17] 기업의 대시보드 구성 1

먼저, 기업에 대한 소개와 기업의 업종, 종가, 기업의 가치지표를 확인할 수 있다.

[그림 18] 기업의 대시보드 구성 2

**6.1 대시보드 - 기업소개**

[그림18] 좌측에서 확인할 수 있으며, 기업소개는 FnGuide에서 수집한 데이터를 기반으로 제공한다. 데이터베이스에 적재한 기업정보를 활용하여 기업에 대한 정보를 제공한다.

**6.2 대시보드 - 기업의 재무데이터 제공**

[그림18] 우측에서 확인할 수 있으며, DB에 적재한 재무데이터를 기반으로, 간단한 재무 정보 및 기업의 업종을 제공한다.

**6.3 대시보드 – 일봉데이터 제공**

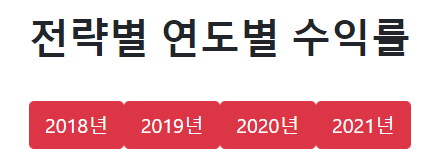
plotly 라이브러리를 활용하여, 종목의 일봉 데이터를 제공한다. 시각화에 활용하는 데이터는 주가의 시가, 고가, 저가, 종가 데이터로, 종목에 대한 주가데이터를 활용하여 파이썬 plotly를 활용하여 동적인 시각화 그래프를 제공하며, 해당 그래프를 json으로 저장 한다. 저장된 json을 웹을 통해 제공한다. html을 사용하여도 무방하지만, plotly 라이브러리의 특성상 html을 사용하는 것보다 웹을 활용하여 제공하기 용이하므로 json을 사용하였다.

**6.4 대시보드 – 관련뉴스**

구글 뉴스를 이용하여 해당 종목의 뉴스를 화면크기 제약으로 인해 최신 뉴스 3개만 불러오도록 구현하였다.

**7. 전략별 백테스트 시각화 차트 제공**

[그림 18] 전략별 백테스트 시각화 제공

전략별로 추천된 종목을 담고 있는 [그림18] 우측 하단의 “종목 전체 연도별 수익률->” 버튼을 클릭하게 되면, 전략을 통해 추천된 과거의 종목을 확인할 수 있다. 또한, 과거 추천된 종목을 모두 매수하였을 경우, 그때 추천된 종목을 1년간 보유하고 있었을 경우의 수익률을 확인할 수 있다

.

위 [그림19]에서 각 연도를 누르게 되면, 해당 전략의 연도별 추천 종목과, 백테스트 기반의 수익률을 확인할 수 있다. 해당 백테스트의 결과는 위의 3.4에서 다룬 결과와 동일한 시각화를 제공한다. 3가지 전략에 대해서도 진행 가능하고, 연도별로 수익률을 확인 가능하다.

2021년의 경우에는 [그림18]에 추천된 종목을 기반으로의 백테스트를 진행한 것이다.



위 그림처럼 해당 년도를 클릭 시 그해 전략으로 사용시 수익률이 제일 좋았던 상위 종목 3개만 추려 보여지도록 하였다.

**8. 배포기능**

Travis CI 를 이용하여 깃허브로 파일을 커밋하면 AWS로 자동으로 배포할 수 있도록 하였다.

**4. 추가내용**

**4.1 한계점**

- **Rule-base기반의 추천**

본 설계에서의 한계는 Rule-base기반의 재무데이터 기반의 추천이라는 점에 있다. 변수가 다양한 만큼 머신러닝 기법을 활용하여 성능지표를 활용한 종목추천이 가능하다고 생각하였지만, 기업들의 재무데이터가 결측치값을 갖는 경우도 존재하였으며, 업종에 따라서 전략에 활용될 지표 데이터가 존재하지 않는 경우도 있다

**- 재무데이터만을 활용한 추천**

본 설계에서는 재무데이터만을 활용한 추천을 진행한다. 현재 가격은 고려되지 않은 채, 재무 상태가 좋은 데이터만을 활용한다는 한계를 갖고 있다. 아무리 재무상태가 좋은 종목이라도 현재 주가가 비싸다면 수익을 내기에는 어려울 것이다. 보조지표를 활용하여 재무 데이터 기반으로 추천된 종목을 싼 가격에 매수하고, 비싸게 매도를 한다면 더 좋은 수익률이 나올 것이다. 하지만, 보조지표를 활용한 간단한 전략을 구성하여, python backtest 라이브러리를 활용하여 간단한 시뮬레이션을 진행해봤을 때, 오히려 더 수익률이 좋지 않아서 구현에서는 제외하였다. 사용한 전략으로는 초기 설계서에 넣은 지지선 돌파, 골든크로스 등의 전략이 있다. 매수시점까지 파악하는 기능이 추가된다면 더 좋은 성능을 보일 것이다.

데이터가 2017~2020년도로 한정되었다는 단점이 존재한다. 분기별 데이터를 활용한다면 3개월을 주기로 투자를 진행해야 하는데, 재무데이터를 활용하여 투자를 하는 것은 보통 가치투자로 일컫을 수 있으며 최소 1년이상 투자를 해야 그 성과가 나온다고 판단하였기에 1년이라는 주기를 갖는 데이터를 활용하여 매년 12월의 재무데이터만을 활용하였다. 활용한 종목에서 더 과거의 데이터인 2016년 이전의 데이터가 존재하였다면 전략에 대한 신뢰성이 더 높아졌을 것이다.

**4.2 개선사항**

**- 전략적 개선사항**

본 프로젝트에서 초기 설계 단계에서 기술적 분석까지 활용하고자 하였다. 하지만, 지지선 산출과 보조지표에 대한 묘둘을 구현하였다. 설계서에서는 MACD, 이동평균선, RSI를 활용하여 매수 매도 구간을 추천해주는 과정을 구현하고자 하였다. 아무리 좋은 종목이라도, 비싼 가격에 매수를 한다면 손해를 볼 수 밖에 없기 때문이다. 본 프로젝트에서는 재무데이터에 기반한 기본적분석(Fundamental Analysis)만을 활용하였지만, 가장 필요한 개선사항으로써 보조지표를 활용한 기술적 분석을 통해 더 좋은 종목을 추천해줄 수 있다. 또한, 전략을 구성하는데에 있어서 구현한 전략이 제한적이다. 다양한 전략이 존재할 수 있지만, 수익률이 좋지 않은 전략도 존재하였고, 검증되지 않은 전략도 많았다. 더 좋은 전략을 구성하고자 다양한 재무데이터를 활용할 수 있다. 또한, 머신러닝을 활용하여 다양한 변수를 활용함으로 써 더 좋은 성능을 보일 수 있으며, 성능평가 지표를 활용하여 눈에 띄는 검증이 가능할 수 있다.

딥러닝 RNN 기반의 LSTM 모델 혹은 ARIMA와 같은 시계열 모델을 통하여 주가를 예측하도록 구현할 수 있다. 초기 구상으로는 1년후의 특정 이동평균선(20일선)을 Label값으로 하고, 거래량과 현재 주가, 보조지표들을 변수로 하여 학습을 진행하여 주가를 예측하고자 하는 방향으로도 진행하였다. 하지만, 본 프로트는 단순히 지수를 예측하거나 한 두 종목을 대상으로 하는 것이 아닌 783개의 종목을 대상으로 진행하여, 모델링을 하는데에 있어서 어려움을 겪었다. 783개의 종목에 대해서 서로 다른 모델을 통하여 학습 및 예측을 진행하는 것은 서비스를 운영하는데에 있어서 어려움이 있을거라 생각했다. 주가 예측과 관련된 다양한 논문들을 참고해봤지만, 대부분 지수를 예측하거나, 한 두 종목에 대한 예측에 불과하여 본 주제에 적용하기에 어렵다고 판단하였다. 대시보드에서 주가 예측기능을 제공함으로써, 종목을 매수하는 데에 있어서 많은 도움을 줄 수 있또록 개선할 수 있다고 생각한다.

또한, 기술적 분석으로도 골든크로스, 군집화를 활용한 지지선 산출 및, 지지선 돌파구간 매수, RSI 등 다양한 지표를 변수로 하여 다양한 전략을 구성하여 진행하고자 하였지만, 전략을 구성하여 백테스트를 한 결과 수익률이 많이 저조하였다. 기술적 분석 전략을 구성하는데에 있어서 많은 어려움을 겪어, 이 점 또한 큰 개선사항이다.

**- 기술적 개선사항**

사용자에게 퀀트투자라는 개념은 흔하지 않아 단순한 종목 추천을 한다해도 믿을만한 베이스 자료들을 사용자가 또 공부해야 한다는 단점이 존재한다. 또한 현재 사용한 보조지표 이외에도 수많은 보조지표들이 존재한다. 그걸 공부하고 사용자가 직접 보기엔 쉽지 않다. 투자 심리선, 갭상승 등 다양한 상황을 고려할 수 있는 보조지표를 기반 전략들에게 같이 구현할 수 있었다면 사용자에게 더욱 더 좋은 기반 데이터를 제공할 수 있었을 것이다. 또한 메일링 서비스를 통해 증권사문 리포트를 제공하려 베이스 데이터를 크롤링하여 데이터화 시켰지만 개발 당시 포트 오류,보안 오류 및 템플릿이 화면에서 깨져 발송되어 기능화 시키지 않았지만 기반 데이터와 템플릿을 따로 만들어 제공 할 수 있었다면 사용자에게 좀 더 친화적인 서비스를 제공 할 수 있었을 것이다. 뿐만 아니라 관심종목 설정은 사용자 별로 따로 구분 값이 필요한 관계로 구글, 네이버 로그인 서비스와 같이 기능 추가적 개선사항이다.

**4.3 참고문헌**

[1] Baker, Malcolm, Brendan Bradley, and Ryan Taliaferro. 2014. “The Low-Risk Anomaly: A Decomposition into Micro and Macro Effects.” Financial Analysts Journal 70

[2]Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2009. “High Idiosyncratic Volatility and Low Returns: International and Further US Evidence.” Journal of Financial Economics 91

[3] 퀀트로 가치투자하라 [웨슬리 그레이, 토비아스 칼라일 저/서태준 역/신진오 감수 | 에프엔미디어]

[4] Greenblatt, Joel. 2010. The Little Book That Still Beats the Market. Vol. 29. John Wiley & Sons. Novy-Marx, Robert. 2013. “The Other Side of Value: The Gross Profitability Premium.” Journal of

[5] Piotroski, Joseph D, and others. 2000. “Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers.” Journal of Accounting Research 38:

[6] An investigation of the relative strength index(RSI)

B Anderson, S Li - Banks & bank systems, 2015

[7] <https://securities.miraeasset.com/hki/hki3028/r01.do> 미래에셋 증권용어사전

[8] Stock market efficiency and the **MACD**. Evidence from countries around the world - Gabriel Dan I.Anghel

**동영상링크**

**https://www.youtube.com/watch?v=G\_ZSEanFxPA**